

Manuel Fernando Moreira da Silva

**Localização de Ontologias de Domínio
no Contexto de Redes Colaborativas**

Julho, 2012

Tese apresentada para cumprimento dos requisitos necessários à obtenção do grau de
Doutor em Linguística, na Especialidade de Lexicologia, Lexicografia e Terminologia,
realizada sob a orientação científica da Professora Doutora Maria Rute Vilhena Costa e sob
a co-orientação do Professor Doutor António Manuel Lucas Soares

Apoio financeiro da FCT no âmbito do PROTEC 2008 - Programa de apoio à formação
avançada de docentes do Ensino Superior Politécnico

SFRH/PROTEC/49350/2008

ao meu filho Rodrigo

AGRADECIMENTOS

Ao terminar a minha dissertação não posso deixar de agradecer a todos os que contribuíram, tanto a nível académico, como pessoal, para que este momento se tornasse possível:

À Professora Doutora Maria Rute Vilhena Costa, minha orientadora, pela sua atitude de exigência, pelos seus preciosos ensinamentos, pelo constante e profícuo debate científico, pela sua disponibilidade e por ter acreditado em mim nos momentos mais difíceis.

Ao Professor Doutor António Manuel Lucas Soares, meu co-orientador, pela valiosa troca de ideias e perspectivas e pelo convite para o INESC Porto, que me proporcionou a oportunidade de crescer científica e pessoalmente.

Aos especialistas do domínio da reabilitação que amavelmente aceitaram participar nesta investigação e partilhar o seu saber e o seu tempo.

Aos colaboradores do PAOL, Dr.^a Célia Tavares e Dr.^a Luciana Oliveira, e do CML, Dr. Pedro Duarte, pela sua inestimável ajuda e disponibilidade.

Às minhas colegas e amigas do ISCAP, Alexandra Albuquerque, Joana Fernandes, Graça Chorão, Maria Clara Cunha, Paula Almeida, Paula Peres, Teresa Pataco, Sara Cerqueira e Suzana Cunha, pela sua amizade, ajuda e incentivo.

Aos colegas investigadores do INESC Porto, pela ajuda e trabalho conjunto, em especial à Doutora Carla Pereira e ao Dr. Cristóvão Sousa.

À minha família, muito especialmente aos meus pais e ao meu irmão, pelo apoio e carinho.

À Sandra, pela dedicação constante e por me manter no rumo.

E, a todos os meus amigos, cujo apoio e presença foi importante para levar a bom termo esta caminhada.

A todos o meu muito obrigado.

LOCALIZAÇÃO DE ONTOLOGIAS DE DOMÍNIO NO CONTEXTO DE REDES COLABORATIVAS

MANUEL FERNANDO MOREIRA DA SILVA

PALAVRAS-CHAVE: Terminologia, ontologias, localização, conceptualização, especificação multilingue de ontologias, representação do conhecimento, redes colaborativas.

KEYWORDS: Terminology, ontologies, localization, conceptualization, multilingual ontology specification, knowledge representation, collaborative networks.

RESUMO

O presente trabalho de investigação apresenta um método para apoiar a especificação multilingue de ontologias de domínio, através do recurso à localização. O método foi desenvolvido para ser aplicado durante o processo de conceptualização de uma ontologia, sendo a sua aplicação enquadrada no contexto de uma rede colaborativa multilingue. Para a sua construção seguimos uma perspectiva assente numa metodologia de análise interlinguística como base para a conceptualização do domínio, tendo como elemento central o uso de metodologias de base terminológica, num enquadramento teórico que privilegia uma abordagem conceptual à representação do conhecimento, mas que considera os contributos resultantes da abordagem linguística.

Observámos, para a sua construção, as lacunas patentes no acesso à informação e à comunicação especializada multilingue, tendo reflectido sobre o modo como este cenário representa um problema para as organizações internacionais. Reconhecemos, neste contexto, a relevância assumida pelas ontologias como recurso privilegiado para a representação do conhecimento de um domínio e, tendo assumido que as ontologias são independentes de qualquer língua, defendemos que a criação e o acesso ao conhecimento continuarão a basear-se na língua, devendo manter um carácter multilingue, o que conduz a uma necessidade crescente de existência de recursos multilingues, bem como de serviços de localização que contribuam para ultrapassar as dificuldades de disseminação do conhecimento. Nesse sentido, desenvolvemos uma análise ao processo de localização de

ontologias e a diferentes abordagens e ferramentas que promovem a integração e a gestão dos elementos multilingues de uma ontologia.

A crescente existência de espaços multilingues de partilha de informação - como as redes sociais e as redes colaborativas, que se caracterizam pela heterogeneidade e multiculturalidade - cria a necessidade da existência de novos tipos de sistemas de organização do conhecimento que tornem mais eficiente a partilha de conhecimento. No entanto, verifica-se, ainda, uma falha de comunicação entre as representações formalizadas do conhecimento, como as ontologias, desenvolvidas com o intuito de facilitar a reutilização deste conhecimento e a interoperabilidade na Web, e os utilizadores que recorrem à sua língua (natural) para aceder ao conhecimento. Como forma de ultrapassar esta falha, apontámos a necessidade de explorar e desenvolver mais aprofundadamente as técnicas e ferramentas de localização em adequação a este contexto, em conjugação com a promoção de novas abordagens, de carácter multidisciplinar, com recurso a modelos de conhecimento do mundo inspirados, por exemplo, nas redes sociais, visão que procurámos desenvolver ao longo desta investigação.

Neste contexto analisámos o papel desempenhado pela terminologia na representação do conhecimento, mais especificamente, no desenvolvimento dos processos de conceptualização e especificação multilingue de ontologias. Analisámos, para tal, o processo de conceptualização do conhecimento de um domínio, central à construção de uma ontologia, e a interligação decorrente do recurso às práticas e metodologias terminológicas e o eficiente desenvolvimento daquele processo, sobretudo numa fase de especificação semiformal da conceptualização, pelo recurso à representação do conhecimento através de mapas conceptuais. Reflectimos, finalmente, sobre o modo como esta forma de representação sustenta a participação dos especialistas no processo de elicitação e organização do conhecimento e promove o trabalho interdisciplinar, nomeadamente com o terminólogo.

ABSTRACT

In this study, a method to support the multilingual specification of domain ontologies, through the use of localization, is proposed. The method was developed to be applied during the conceptualization process in the context of a collaborative network and is based on an interlinguistic analysis that functions as support for the conceptualization of the subject field. It is terminology-based, and follows a theoretical framework that recognizes the conceptualization process as the central point for developing knowledge representation in more than one language. Thus, it is underpinned by a conceptual approach to knowledge representation, although it also considers and integrates contributions from the linguistic approach.

For its construction, we have analyzed the existing gaps in the access to specialized communication and information in multilingual environments and reflected on how this scenario poses problems to international organizations. In this context, ontologies arise as an important resource to support knowledge domain representation. Although we recognize ontologies as language independent artifacts, it is our perspective that knowledge creation and access are language-based and thus multilingual, which leads to the need for multilingual resources development, as well as localization services that help overcome the difficulties in the process of knowledge dissemination. To better understand how to support this issue, we have studied the process of ontologies localization and analyzed the different approaches and tools proposed for the integration and management of multilingual elements of an ontology.

The growing existence of multilingual environments that promote knowledge and information sharing – like social or collaborative networks, characterized by heterogeneity and multiculturalism – raises the need for the existence of knowledge management systems that promote a more efficient knowledge sharing. However, we still face a lack of communication between formalized knowledge representations, developed to support interoperability between systems and the reuse of knowledge, and the users who utilize their (natural) language to access knowledge. To lessen this gap, we see a need to further explore and develop localization techniques and tools adequate to this context, in conjunction with the promotion of new multidisciplinary approaches, based on models of world knowledge inspired, among others, in social and collaborative networks, a vision that we have tried to develop throughout this research.

Taking into consideration this context, we paid special attention to the role played by terminology in supporting the knowledge representation process, more specifically on the development of the processes of conceptualization and multilingual specification of ontologies. For this purpose we focused on the process of construction of a conceptualization of the domain's knowledge, central to the construction of an ontology, and its interconnection between the use of terminological practices and methodologies and the efficient development of conceptualization, especially in a phase of semiformal specification of the conceptualization, based on the use of conceptual maps to support the knowledge representation. Finally, a reflection was conducted on the way this representation environment supports and enhances the specialists participation in the elicitation and organization of knowledge processes and how it promotes the interdisciplinary work, namely with terminologists.

ÍNDICE

Introdução	1
1. Multilinguismo: elemento de desenvolvimento e disseminação do conhecimento	15
1.2. Organização e partilha do conhecimento	21
1.2.1. Conceito	24
1.3. Conceptualização	33
1.3.1. Especificação semiformal da conceptualização	36
1.3.2. Terminologia e conceptualização.....	40
1.4. Ontologias	44
1.4.1. Terminologia e ontologias.....	50
1.5. Terminologia e organização do conhecimento	56
1.5.1. Perspectiva linguística	59
1.5.2. Perspectiva conceptual	66
1.6. Representação do conhecimento em contextos multilingues.....	71
2. Localização	79
2.1 Localização e Tradução	82
2.2 Terminologia e Localização	85
2.3 Localização de ontologias	89
2.3.1 Localização de ontologias: dimensões e dificuldades	93
2.4 Abordagens ao processo de localização de ontologias	100
2.4.1 Ferramentas de base linguística.....	105
2.5 Localização e localiação de ontologias: analogias.....	113
3. Método para a especificação multilingue de ontologias	119
3.1. Objectivos e justificação do método.....	121
3.2. Visão geral do método	123

3.3.	Contributo da terminologia	129
3.3.1.	Constituição do <i>corpus</i>	129
3.3.1.1.	Recursos linguísticos e semânticos	131
3.3.1.2.	Corpus Textual	137
3.4.	Papel do terminólogo	142
3.5.	Função da língua inglesa na conceptualização e localização do domínio	145
3.6.	Contributo da representação do conhecimento	146
3.7.	Ferramentas e ambientes de suporte ao desenvolvimento do método	151
3.7.1	Espaço de conceptualização e localização	152
3.7.2	Aplicações e recursos linguísticos de suporte à localização.....	157
3.8.	Fluxo de trabalho do método	161
3.9.	Papel dos especialistas de domínio	163
4.	Especificação multilingue de ontologias: experimentação	173
4.1.	Caso de estudo – especificação multilingue da ontologia H-Know.....	173
4.2.	Implementação do método no contexto do caso de estudo	178
4.2.1.	Fase 1 – Selecção de recursos do domínio e criação do ambiente de localização	182
4.2.1.1.	Domínio de especialidade: caracterização e delimitação	183
4.2.1.2.	Compilação do corpus	188
4.2.2.	Fase 2 – Definição dos participantes e da amostra	198
4.2.2.1.	Definição dos especialistas.....	198
4.2.2.2.	Definição da amostra.....	200
4.2.3.	Fase 3 - Criação do espaço de conceptualização e localização	202
4.2.3.1.	Definição dos recursos e aplicações de apoio à localização.....	205
4.2.4.	Fase 4 – Experimentação do método	209
4.2.4.1.	Desenvolvimento de um guião para o especialista	209
4.2.4.1.1.	Acompanhamento do especialista na aplicação do método	210
4.2.5.	Fase 5 – Análise de resultados	211

4.2.5.1.	Análise dos resultados da experimentação do método.....	212
4.2.5.1.1.	Desenvolvimento do processo de localização	214
4.2.5.1.2.	Desenvolvimento do processo de (re)conceptualização	221
4.2.5.2.	Análise prévia dos resultados.....	234
4.2.5.3.	Inquérito aos especialistas	237
4.2.6.	Fase 6 - Reestruturação do método.....	244
	Conclusão	249
	Bibliografia	267
	ÍNDICE DE FIGURAS	295
	ÍNDICE DE TABELAS	297
	APÊNDICES	i
	Apêndice A: Critérios para a avaliação de fontes.....	iii
	Apêndice B: Guião para o especialista	v
	Apêndice C: Guide for the expert.....	x
	Apêndice D: Entrevista aos participantes no caso de estudo	xv
	Apêndice E: Interview to the participants in the case study.....	xix
	Apêndice F: Entrevista - Recolha e análise de resultados.....	xxiii
	Apêndice G: Corpus do domínio da reabilitação	xxxi
	Apêndice H: Mapas conceptuais: amostra.....	xxxii
	ANEXOS	xxxviii
	Anexo 1: Tipologias de recursos.....	xl
	Anexo 2: Comparação entre estruturas organizacionais do processo construtivo	xli
	Anexo 3: Exemplo de Ficha Patorreb	xlii
	Anexo 4: Implementação da ontologia H-Know em Protégé.....	xliii

INTRODUÇÃO

Introdução

We dissect nature along lines laid down by our native languages. The categories and types that we isolate from the world of phenomena we do not find there because they stare every observer in the face; on the contrary, the world is presented in a kaleidoscopic flux of impressions which has to be organized by our minds—and this means largely by the linguistic systems in our minds. We cut nature up, organize it into concepts, and ascribe significances as we do, largely because we are parties to an agreement to organize it in this way - an agreement that holds throughout our speech community and is codified in the patterns of our language. The agreement is, of course, an implicit and unstated one, but its terms are absolutely obligatory; we cannot talk at all except by subscribing to the organization and classification of data which the agreement decrees. (Whorf 1940/1972)

A Terminologia, enquanto disciplina científica, é uma disciplina consagrada ao estudo científico dos conceitos e termos usados nas línguas de especialidade e à sua estruturação em sistemas de representação do conhecimento especializado, o que a torna crucial para o desenvolvimento dos processos de comunicação e a partilha de informação e de conhecimento, que ocupam um lugar preponderante no desenvolvimento e evolução da sociedade moderna globalizada.

Neste contexto, a terminologia disponibiliza um conjunto de metodologias que visam facilitar e promover o acesso ao conhecimento e a sua gestão, concentrando-se o trabalho terminológico na recolha, descrição, processamento e representação sistemática dos conceitos específicos a um domínio de especialidade e das suas designações. Tem, como descreve (Roche 2012: 17), como objectivos principais compreender o mundo, descrever os objectos que o povoam e encontrar as palavras certas para os descrever.

O recurso à terminologia tem, em consequência, crescido em importância e visibilidade à medida que a procura, a produção, o processamento e a disseminação de informação se tornam centrais ao processo de comunicação em termos globais, sobretudo

quando se consideram organizações internacionais, redes e comunidades de profissionais multinacionais, ou se tem em conta fenómenos como o aumento constante do número de terminologias específicas que acompanham a evolução dos diferentes domínios do saber.

Como destaca a Unesco, o nível de desenvolvimento relativo de um país pode ser medido a partir da capacidade média que os seus cidadãos possuem em utilizar a informação com o intuito de transferir conhecimento e reforçar as suas capacidades internas. Nesse sentido, e dado que:

terminological data constitute the core element of any domain-specific information and knowledge representation, the availability of and accessibility to these data are critical socio-economic factors. (UNESCO, 2005: 02)

O desenvolvimento tecnológico constante, por um lado, e o aumento da cooperação a nível nacional e internacional, por outro, levam a que recursos de informação estruturados possam e devam ser reutilizados, integrados e disponibilizados de modo a simplificar e a promover o acesso à informação em diferentes línguas.

O avanço na direcção de uma cultura de gestão da informação verdadeiramente multilingue, resultado do estabelecimento de espaços multilingues de acesso à informação, caracterizados pela heterogeneidade e multiculturalidade, como é o caso dos sistemas de classificação de organizações globais, aumenta a necessidade de existência de metodologias e sistemas de organização da informação e do conhecimento que facilitem e tornem mais eficiente o uso destes recursos.

Aos fenómenos anteriores acresce o fenómeno da globalização, o aumento da popularidade da Internet e a disponibilidade, a nível global, de fontes de informação distribuídas em rede, cuja existência conduziu ao aumento da procura de informação multilingue e à necessidade de desenvolver tecnologias que apoiassem os processos de aquisição, modelação, navegação, representação, visualização e acesso à informação e ao conhecimento por membros oriundos de comunidades linguística e culturalmente diversas.

Por outro lado, a análise da informação existente em cada organização e disseminada no seu contexto de acção, torna-se essencial ao desenvolvimento de novos processos de gestão, projectos, serviços e produtos e à articulação integrada entre todos estes elementos. Esta necessidade torna-se ainda mais manifesta quando a análise, a representação e o

acesso ao conhecimento acontecem em domínios profissionais cujo processo produtivo envolve normalmente um elevado número de entidades, grandes exigências ao nível do planeamento e onde o conhecimento se encontre disperso.

Para fazer face às exigências crescentes na complexidade dos processos intra e interorganizacionais, sobretudo quando estes se desenrolam num contexto multinacional, tem-se verificado um aumento de qualidade nos modos de interacção e partilha dos recursos das organizações. A resposta aos desafios colocados por este tipo de contexto tem acontecido através, por um lado, de um maior envolvimento dos diferentes interlocutores em formas eficazes e inovadoras de colaboração, como é o caso da criação de redes colaborativas, ao nível empresarial e, por outro, do desenvolvimento de sistemas de gestão da informação e do conhecimento mais robustos, como os sistemas de gestão do conhecimento baseados em ontologias.

Estes sistemas têm como fim promover uma maior robustez na pesquisa, acesso e a reutilização da informação e do conhecimento disponibilizado na Web, sobretudo o que respeita a domínios técnicos altamente especializados. Contudo, as abordagens correntes ao à disponibilização e acesso à informação em mais do que uma língua oferecem apenas soluções parciais que encaram o problema de modo restrito, pelo que a urgência em aceder ao conhecimento, ultrapassando as barreiras impostas pelas diferentes línguas naturais, se tornou uma necessidade e um desafio premente para o desenvolvimento e uso eficaz dos diferentes sistemas de gestão do conhecimento.

Partindo de uma análise da literatura existente, apercebemo-nos de que a evolução da preocupação com os fenómenos multilingues e a inclusão de abordagens que considerem o multilinguismo no desenvolvimento de ontologias esteve entre os primeiros problemas e dificuldades detectados na evolução da engenharia de ontologias. No entanto, a integração e o enriquecimento das ontologias com elementos multilingues foi sendo investigada de modo menos sistemático, dadas as muitas dificuldades associadas à construção de ontologias com recurso a mais de uma língua.

Só mais recentemente se notou uma evolução nos sistemas técnicos e abordagens teóricas disponíveis, bem como na evolução de alguns princípios epistemológicos que possibilitaram a abertura de novos campos de reflexão e análise nesta área. No entanto, uma vez que o conhecimento e as representações do conhecimento não se restringem ao

uso de uma língua em particular, o enriquecimento dos sistemas de gestão da informação com informação multilingue tornou-se uma necessidade evidente, sobretudo na construção de ontologias, fenómeno que pudemos constatar ao longo da nossa participação em projectos na área da gestão do conhecimento e cuja observação se tornou o ponto de partida para o desenvolvimento da problemática e das reflexões que apresentamos nesta tese.

O número diminuto de ontologias formalizadas que contenham, associada, informação linguística em mais do que uma língua provém de vários factores, a começar pela dificuldade inerente à escolha das metodologias a adoptar para a conceptualização e representação do conhecimento num ambiente de desenvolvimento e localização de ontologias para diferentes línguas e pela falta de ferramentas específicas que apoiem este processo, ferramentas que produzam resultados de maior qualidade e cujo uso considere utilizadores sem formação específica em tradução ou localização.

Outras dificuldades advêm, sobretudo, de problemas de ordem linguística, que surgem, por exemplo, na associação dos significados de termos de diferentes línguas a conceitos, uma vez que os termos usados para designar um determinado conceito no sistema conceptual podem não encontrar correspondente quer nas representações textuais do domínio quer nos diferentes sistemas de classificação e taxonomias.

De modo a responder a estas dificuldades, a abordagem tradicionalmente seguida na construção de ontologias multilingues é desenvolvida recorrendo a uma primeira representação conceptual monolingue, que pode ser feita a partir de uma língua convencionada, geralmente o Inglês, sendo a localização apenas realizada na fase final do processo da sua construção, após a sua formalização. Este processo é desenvolvido com recurso à localização e acontece na fase seguinte ao da especificação formal da ontologia, sendo visto sobretudo como uma questão de enriquecimento da ontologia com elementos multilingues, questão a resolver após a construção e formalização da ontologia.

Na nossa opinião, a elaboração de uma metodologia que contribua para que a localização conduza a uma melhor especificação multilingue de uma ontologia e, conseqüentemente, à sua mais efectiva adaptação a uma ou mais comunidades linguísticas e culturais concretas, e que, em paralelo, contribua para potenciar a interoperabilidade semântica entre diferentes recursos no âmbito de um domínio específico, requer que

tenhamos em conta o contexto de construção da ontologia e que nos focalizemos na apreensão da complexidade, riqueza e diversidade do domínio e na sua representação em termos linguísticos na fase de conceptualização da ontologia. Requer, também, que se tenham presentes e considerem os objectivos e os propósitos de uso da ontologia pela comunidade de especialistas que procede à sua representação, bem como pela comunidade de utilizadores potenciais.

O objectivo mais genérico do desenvolvimento de processos de localização de ontologias é o de permitir a interoperabilidade interlinguística em sistemas de informação que incluem, normalmente, um grande número de recursos de conhecimento. A construção de uma abordagem que potencie a interoperabilidade semântica entre estes recursos requer, por isso, uma análise e modelação detalhadas dos contextos profissionais e colaborativos onde o conhecimento é transmitido e estruturado, e a definição de um método que nos permita capturar e representar o conhecimento do domínio, tendo por base o conhecimento do especialista sobre o domínio e a informação existente nos diferentes tipos de recursos disponíveis, como *corpora*, classificações, terminologias, tesouros, entre outros.

Tendo em conta os aspectos descritos, é nosso objectivo apresentar nesta dissertação um método para apoiar a especificação multilingue de ontologias de domínio, através do recurso à localização. Por especificação multilingue de uma ontologia entendemos as diferentes fases que compõem um processo de localização de uma ontologia desenvolvido no momento consecutivo à sua conceptualização e anterior à sua especificação formal.

Este método foi concebido de modo a ser aplicado durante o processo de conceptualização de uma ontologia, no momento subsequente a um processo de conceptualização inicial, e foi desenvolvido tendo em perspectiva a sua aplicação e experimentação no contexto de uma rede colaborativa multilingue.

O método tem como objectivos propor um espaço de trabalho colaborativo para apoiar especialistas e terminólogos na localização de ontologias para diferentes línguas no momento da sua conceptualização e permitir a integração dos diversos contributos resultantes da interacção entre os diferentes interlocutores envolvidos na conceptualização.

Pretende, finalmente, contribuir para o processo de análise e representação do conhecimento do domínio no contexto de redes colaborativas multilingues.

Pretendemos, através da implementação deste método, criar condições para que no processo de localização se escolham os equivalentes apropriados tendo por base o conceito, *i.e.*, que a preocupação dos diferentes interlocutores envolvidos no processo de localização não recaia somente em encontrar o equivalente mais correcto do ponto de vista da língua, mas antes em encontrar o equivalente mais correcto em termos conceptuais, de modo a permitir uma representação mais efectiva do conhecimento do domínio nas diferentes línguas e, posteriormente, a sua mais fácil integração na ontologia de domínio representada em mais do que uma língua.

Assumimos assim, como ponto de partida para o desenvolvimento do processo de especificação multilingue da ontologia, o uso de uma metodologia de base terminológica. Esta metodologia adoptará uma perspectiva conceptual, tendo como ponto de partida o conceito, e será complementada, sempre que se justifique, com os contributos obtidos pelo recurso à perspectiva linguística.

Para a construção deste método considerámos um conjunto de três hipóteses de partida. Assim, para a construção deste estudo, a nossa observação parte da análise do processo de conceptualização do conhecimento de um domínio e da construção de uma ontologia de domínio representada em mais do que uma língua natural no contexto de uma rede colaborativa multilingue. Neste contexto, a representação do conhecimento acontece com recurso negociação e obtenção de consenso entre os especialistas do domínio, num processo onde as culturas, as experiências e os recursos de cada actor/especialista estão presentes, criando um espaço de conceptualização mediado por uma intercultura profissional.

Este contexto e os aspectos supramencionados levam-nos a colocar uma primeira hipótese: a de que a localização, ao ser realizada numa fase subsequente à da conceptualização e anterior à da especificação formal de uma ontologia, através do uso de mapas conceptuais, pode representar um incremento efectivo à construção de uma conceptualização em mais do que uma língua e enriquecer o processo de partilha do conhecimento durante a construção de uma ontologia de domínio ao:

1. permitir que se desenvolvam representações conceptuais do conhecimento do domínio utilizando, para a sua descrição, as línguas de trabalho dos diferentes especialistas da rede;
2. permitir um *feedback* mais eficaz sobre a conceptualização do domínio a partir das representações semiformais de conhecimento obtidas com recurso a mapas conceptuais;
3. promover uma maior participação dos especialistas na construção e na localização da ontologia, tornando a sua implementação mais exequível, na medida em que poderá conduzir a um envolvimento mais efectivo do especialista nos diferentes momentos da sua construção e, eventualmente, conduzir a um processo de reconceptualização, ao levar a que se activem e considerem os pontos de vista e as características específicas às línguas e culturas em presença.

Quando consideramos a especificação multilingue de ontologias, desenvolvida com recurso a um processo de localização, situamo-nos perante a localização de um artefacto que é, em si mesmo, uma representação estruturada de conhecimento específico a um domínio, pelo que nos confrontamos com um objecto que deve ser entendido como um todo. Acreditamos que numa ontologia o significado de uma das suas partes, de um dos seus conceitos, não é autónomo, mas depende dos outros conceitos com que se relaciona e que o significado global da ontologia não é o resultado de uma mera soma das suas partes, mas de uma combinação geradora de sentidos.

Atendendo a esta perspectiva e ao momento em que propomos que a localização se desenrole – consecutivamente a uma conceptualização inicial -, considerámos uma segunda hipótese: a de que a localização representa, ela própria, enquanto processo, um elemento que contribui para a construção mais efectiva de uma conceptualização do conhecimento do domínio, facilitando, posteriormente, os processos de especificação formal e de manutenção da ontologia.

Esta hipótese surge porque acreditamos que o processo de localização pode, por via do discurso, contribuir para explicitar de forma mais precisa o conhecimento, através, por exemplo, do modo como o especialista reflecte sobre a conceptualização inicial, da forma como coloca as dúvidas à comunidade ao longo do processo de escolha dos equivalentes ou da forma como constrói a sua lógica discursiva.

Ao recentrarmos o processo de localização, situando-o no momento subsequente ao do processo de conceptualização, e ao desenvolver o processo a partir da representação do conhecimento sob a forma de mapas conceptuais, afastamo-nos dos processos mais correntes de localização dos termos explicitados formalmente numa ontologia e instituímos os conceitos, e os termos que os designam, representados em mapas conceptuais, como elementos a localizar.

O processo de localização permite, neste contexto, que se considerem não só os elementos que constituem o sistema conceptual – conceitos e relações - como acontece quando a localização ocorre a partir de uma ontologia formal, mas antes, e sobretudo, que os actores envolvidos no processo possam ter em conta a representação conceptual do conhecimento do domínio como um todo, a partir da sua visualização através de mapas conceptuais e uma percepção mais contextualizada da interligação entre os conceitos.

O lugar central que atribuímos ao processo de conceptualização leva-nos a colocar uma terceira hipótese: a de que o processo de conceptualização de uma ontologia de domínio, desenvolvido de forma partilhada no contexto de uma rede colaborativa multilingue pode condicionar positivamente o processo de localização, tal como é tradicionalmente percebido. De facto, deixamos de estar perante a localização de termos isolados ou descontextualizados, feita por um tradutor ou pelo recurso a sistemas de tradução automática. Partimos de uma representação conceptual inicial do conhecimento do domínio, construída por especialistas de domínio, eventualmente com a participação de terminólogos e gestores de conhecimento, para a sua localização, com o contributo do conjunto dos interlocutores.

De modo a testar estas hipóteses e a proceder à experimentação do método consideramos como necessário seguir uma abordagem interlinguística, de cariz terminológico, em que o percurso parta da representação do conceito para a pesquisa dos equivalentes apropriados. Propomos esta abordagem dado que acreditamos que a comunicação de especialidade – mono e multilingue – não é apenas um problema de língua ou uma simples questão de equivalência de denominações, mas também de conhecimento: conhecimento dos conceitos que constituem um domínio, das relações que estabelecem entre si e do modo como esse conhecimento é expresso pela língua.

Daremos, assim, atenção à relação entre a construção de uma ontologia e o seu enriquecimento com elementos linguísticos e terminológicos multilingues, com o fim de perceber o modo como a integração de elementos multilingues pode ser melhorada no processo de construção de ontologias representadas em mais do que uma língua natural, no contexto de uma rede colaborativa. Para tal teremos em conta as questões que se colocam na conceptualização de uma ontologia, do ponto de vista conceptual e do ponto de vista linguístico, tendo em conta aspectos como:

- o processo de interpretação e representação estruturada do conhecimento num domínio específico;
- as necessidades e dificuldades do processamento multilingue da informação, sobretudo quando levado a cabo por especialistas de domínio sem formação em tradução ou localização;
- a reutilização de recursos semânticos previamente estruturados, como classificações, tesouros ou ontologias e a sua integração no processo de conceptualização e localização.

Sendo o processo de localização de ontologias, no nosso caso, um processo contextualizado socialmente no seio de uma rede colaborativa, atentaremos ao modo como se desenrola este processo enquanto evento situado cultural, profissional e socialmente.

A análise das questões, a construção do método e o estudo dos processos centrais à nossa dissertação – o de conceptualização de uma ontologia e o da sua especificação multilingue – terão por base uma abordagem de base terminológica, que partirá de uma perspectiva conceptual, de base onomasiológica, i.e., do conceito para o termo, sendo a primazia do conceito sustentada por elementos conceptuais, resultantes, no nosso caso, essencialmente, do conhecimento transmitido pelos especialistas do domínio, perspectiva que será conjugada e complementada com o recurso a uma perspectiva linguística de base semasiológica.

Iniciámos a dissertação com um primeiro capítulo em que reflectimos acerca do multilinguismo e da sua importância enquanto elemento promotor e disseminador do conhecimento, sobretudo o especializado. Para tal analisamos os conceitos e ambientes subjacentes à organização e partilha do conhecimento, bem como os conceitos e processos centrais a esta organização – conceito e conceptualização -. Analisámos, para tal, algumas

questões relacionadas com o conceito e com o estatuto do termo, equacionando diferentes perspectivas; e o conceito de conceptualização à luz do desenvolvimento de sistemas de gestão da informação e do papel a desempenhar pelos actores de uma comunidade na sua construção, bem como da relação entre as práticas e metodologias terminológicas e o desenvolvimento do processo de conceptualização e construção de ontologias.

Abordamos, em seguida, o conceito de ontologia, equacionando o papel e contributo que o recurso às práticas e metodologias terminológicas oferecem à construção de ontologias. Consideramos, também, o modo como a terminologia, enquanto ciência, aborda a organização e a representação do conhecimento de um domínio de especialidade e descrevemos algumas orientações propostas para esse fim pela perspectiva conceptual em comparação com as orientações propostas pela perspectiva linguística. Debates também, de forma breve, a relação entre o texto e a ontologia. Discutimos, finalmente, a problemática da representação do conhecimento num contexto multilingue.

O segundo capítulo é dedicado à análise do conceito de localização. Estabelecemos, para tal, as principais diferenças entre o conceito de tradução e o de localização e descrevemos o processo de localização, tendo em conta uma abordagem terminológica. Após esta reflexão, centrámos a nossa atenção no conceito e no processo de localização de ontologias, e analisamos as diferentes dimensões e dificuldades que enformam o processo; identificámos e descrevemos algumas das abordagens existentes; identificamos os modelos de base linguística desenvolvidos especificamente para a localização de ontologias e consideramos os seus pressupostos e limitações; identificámos e caracterizámos algumas das ferramentas de localização de ontologias, e fazemos uma reflexão final sobre as analogias passíveis de serem estabelecidas entre os processos de localização e de localização de ontologias.

O capítulo seguinte tem por objectivo apresentar o método que desenvolvemos de apoio à especificação multilingue de ontologias de domínio, bem como a sua sustentação teórica. Neste capítulo apresentamos uma visão geral do método, descrevemos os seus elementos constituintes, as suas fases e os objectivos que nos propomos alcançar com o seu uso. Para tal, descrevemos o conjunto dos contributos advindos da área da terminologia e da área da representação do conhecimento que considerámos para a sua construção.

Consideramos, finalmente, o papel e contributos dos principais actores do processo: o especialista e o terminólogo.

No último capítulo apresentamos a experiência levada a cabo para implementar e validar o método proposto. Começámos, assim, por descrever o caso de estudo e o contexto que sustentou a experimentação do método para a especificação multilingue de ontologias. Descrevemos, em seguida, as suas diversas fases e as actividades desenvolvidas em cada fase para permitir a realização da experiência e a recolha e análise de resultados. Apresentámos, após a análise do desenvolvimento do processo, algumas conclusões prévias sobre a aplicação do método. Estas conclusões são, em seguida, confrontadas com a análise da opinião dos participantes na experiência, obtida através da realização de uma entrevista, em que foi nosso objectivo perceber como entendiam o contributo e validade do método proposto, no que respeita ao apoio à estruturação e representação do conhecimento do domínio, ao apoio ao desenvolvimento do processo de localização, e ao apoio à partilha de conhecimento e obtenção de consenso no contexto de uma rede colaborativa multilingue.

Através da reflexão teórica e metodológica e da experimentação que apresentamos neste trabalho de investigação, esperamos contribuir para a evolução e para a introdução de novas perspectivas sobre o processo de especificação multilingue de ontologias, no que concerne, sobretudo, ao desenvolvimento do processo de localização de ontologias no contexto de uma rede colaborativa.

CAPÍTULO I

1. Multilinguismo: elemento de desenvolvimento e disseminação do conhecimento

Their diversity (of languages) is a diversity not of sounds and signs but of ways of looking at the world. (Kerényi, 1976: xxxi)

Num mundo cada vez mais ligado a nível global, o diálogo intercultural que acontece através da comunicação multilingue é um pré-requisito fundamental para a reflexão e a consciência crítica das identidades locais, constituindo este diálogo o fundamento que permite ultrapassar comportamentos particulares e promove a difusão da equidade social, no âmbito das relações socioeconómicas entre povos (Cortese, 2007: 203).

Este diálogo tornou-se ainda mais premente com o advento da globalização, fenómeno que, apoiado pela rápida propagação da Internet, aumentou a necessidade das organizações acederem, gerirem e disseminarem a informação e o conhecimento num número crescente de línguas. Como realça Bergamashi (2003), o mundo está a encolher à medida que o avanço da comunicação e da computação se desenrolam a um ritmo impressionante. No entanto, e numa afirmação na qual nos revemos, a autora deixa claro não haver qualquer dúvida sobre o facto de que os seres humanos continuarão ligados às suas culturas e línguas de origem, dado que as línguas individuais, e o modo peculiar como expressam a informação e designam os conceitos, representam uma base de conhecimento cultural que não deve ser subestimada.

Porém, o acesso à informação e à comunicação multilingue em termos globais não avançou ao mesmo ritmo que as trocas comerciais, apesar do progresso na criação e uso de recursos multilingues de apoio a este processo e à sua constante expansão. Tal acontece sobretudo devido às dificuldades inerentes ao uso das diferentes línguas para a criação e posterior gestão de conteúdos multilingues.

Este cenário, à medida que o comércio electrónico e as relações interorganizações se vão desenvolvendo no âmbito da internet, pode representar uma lacuna para as organizações e redes internacionais que lidam diariamente com informação multilingue, pelo que se torna essencial suprir estas lacunas de modo a fazer face, por exemplo, ao crescente desenvolvimento de ambientes colaborativos virtuais, como é o caso das redes de

cooperação ou das redes de organizações internacionais, caracterizadas pela multiculturalidade e pelo multilinguismo.

Para responder a estes desafios surgiram, na última década, novos paradigmas com origem na análise dos ambientes de trabalho baseados na *Web* centrados na comunicação, colaboração e co-criação de diversos recursos linguísticos. Apesar deste esforço, um conjunto significativo de barreiras linguísticas, bem identificadas, permanece ainda sem resolução, sobretudo quando o enfoque recai sobre a comunicação multilingue em domínios específicos do saber ou sobre a representação do conhecimento em mais do que uma língua, barreiras que surgem associadas não só à diversidade das línguas e às suas características distintas e complexas, mas também à heterogeneidade, à diversidade cultural ou à diversidade situacional (Cabré, 2004b: 98)¹.

Esta problemática, a que dedicaremos a nossa atenção ao longo do capítulo seguinte, estende-se à construção dos recursos linguísticos que representem o conhecimento em mais do que uma língua, como é o caso dos tesouros e das ontologias ou da construção da Web Semântica, que ainda é grandemente monolingue e usa frequentemente a língua Inglesa como língua de trabalho e de comunicação.

Na visão de Berners Lee *et al.* (2001: 02) a Web Semântica facilitaria o desenvolvimento de métodos automáticos que apoiassem os utilizadores na transmissão e compreensão dos conteúdos produzidos. Para os autores, aquela permitiria a produção de conteúdos acessíveis à máquina que poderiam fornecer

automated translation between the output of a scientific device and the input of a data mining package used in some other discipline (...). These new products will allow users to create relationships that allow communication when the commonality of concept has not (yet) led to a commonality of terms. The semantic web will provide unifying underlying technologies to allow these concepts to be progressively linked into a universal web of knowledge, and will therefore help to break down the walls erected by lack of

¹ Como afirma Briu (2012: 08) *En étudiant, on découvre, avec le temps, que le sens des mots, des signes, des termes et des textes est plus complexe qu'on ne croyait, que les langues sont des systèmes souples, très puissants et que la traduction, qui met en rapport au moins deux langues, est une opération naturelle, possible, mais hyper-complexe et risquée, un défi ou une « négociation ».*

*communication*², and allow researchers to find and understand products from other scientific disciplines.

Esta visão, que expressa sobretudo uma vontade de desenvolvimento, não corresponde ainda à realidade, sobretudo no que se refere à presença do multilinguismo, uma vez que subsiste ainda, a nosso ver, uma falha entre as representações de conhecimento formalizadas, como é o caso das ontologias, e os utilizadores dos sistemas de informação e comunicação, onde essas representações são utilizadas.

Por outro lado, e como referem Benjamins *et al.* (2002), o multilinguismo surge, efectivamente, como um dos desafios à realização da Web Semântica, sendo necessário ter em conta diferentes aspectos que contribuem para que esse desafio seja ultrapassado, de que destacam:

1. *At the ontology level, ontology builders may want to use their native language for the development of the ontologies in which annotations will be based.*³
2. *At the annotation level, annotation of content can be performed in various languages.*
3. *Finally, at the user interface level, millions of people would like to access relevant content in their native language irrespective of the source language in which annotations are presented.*

Nas últimas décadas e em diversas áreas, como a inteligência artificial, as ciências da computação e a terminologia, as ontologias foram ganhando importância enquanto recurso privilegiado para a representação formal do conhecimento de um domínio específico. O objectivo principal de uma ontologia é, como analisaremos com mais pormenor no ponto 1.3, a especificação e representação formal de um conhecimento de modo a assegurar uma descrição mais sistemática e granular deste e um acesso mais eficiente ao mesmo, bem como permitir a sua pesquisa, gestão e reutilização.

Apesar de assumirmos que as ontologias são independentes de qualquer língua, entendemos que a criação e o acesso ao conhecimento far-se-ão por via da língua, e deverão manter, conseqüentemente, um carácter multilingue.

² Nosso sublinhado.

³ Os recursos linguísticos e outros recursos multilingues disponíveis (como o WordNet, EuroWordnet, etc.) podem ser explorados para apoiar o multilinguismo a este nível.

A urgência em aceder ao conhecimento, ultrapassando as barreiras impostas pelas diferentes línguas, tornou-se, nesse sentido, uma necessidade premente. Como reconhece Budin (2009), também no caso do desenvolvimento da Web Semântica se verifica uma necessidade crescente de existência de recursos ontológicos multilingues, que possam ultrapassar as barreiras de comunicação que se erguem a partir das diferenças linguísticas e culturais, da falta de fluência no uso da língua inglesa e da necessidade de rigor na comunicação.

É também amplamente reconhecido que estes recursos são essenciais quer para aceder à informação a partir de diversas línguas, quer para permitir que esta represente e abranja a diversidade cultural transmitida pelas diferentes línguas, pelo que qualquer abordagem ao seu desenvolvimento deve, necessariamente, incluir a possibilidade de aceder à informação em diversas línguas.

No entanto, ainda se verifica uma falha de comunicação entre as representações formalizadas do conhecimento, como as ontologias, desenvolvidas com o intuito de facilitar a reutilização deste conhecimento e a interoperabilidade na Web, e os utilizadores dos sistemas de informação que recorrem à sua língua para aceder ao conhecimento. Na tentativa de ultrapassar essa falha iremos explorar e desenvolver mais aprofundadamente as técnicas de localização em adequação a estas necessidades e a este contexto, de modo a tornar possível o acesso personalizado à informação a partir da língua de origem de cada utilizador.

Por outro lado, as barreiras acima identificadas poderão, a nosso ver, ser mais facilmente transpostas se, se promoverem novas abordagens, de carácter multidisciplinar, com recurso a modelos de conhecimento do mundo inspirados, por exemplo, nas redes sociais, visão que procuraremos ter em conta ao longo deste estudo.

A realidade acima descrita tem pressionado e conduzido ao desenvolvimento de uma variedade de recursos multilingues. O desenvolvimento destes recursos é um processo complexo, em que se identificam vários problemas e dificuldades transversais tanto à construção de recursos terminológicos, como à construção de outros recursos linguísticos.

Estes recursos são recursos previamente estruturados, como classificações, tesouros, *wordnets*, taxonomias ou ontologias, entre outros, nos quais a informação relativa ao

domínio de especialidade já se encontra organizada e estruturada de acordo com as especificidades inerentes a cada tipo de recurso e em consonância com os propósitos particulares que presidiram à sua construção, podendo assumir um formato mono ou multilingue.

Para que estes recursos venham a ter um desenvolvimento mais acentuado, reconhecemos, tal como Calzolari & Soria (2005: 74) que o processamento efectivo de conteúdo multilingue requer uma mudança de paradigma e o desenho de uma nova geração de recursos, que a autora antevê venham a ser baseados em normas interoperáveis de conteúdos abertos.

Calzolari & Soria (2005: 74) chamam, por outro lado, a atenção para a necessidade de cooperação na construção e manutenção destes recursos, ponto de vista com o qual concordamos, ao afirmar que o esforço colocado na disponibilização de milhões de “palavras” para uma multiplicidade de línguas é algo que não cabe no seio do trabalho de um grupo único ou isolado. Este objectivo só se atingirá, na opinião dos autores, se se caminhar na direcção de uma infra-estrutura linguística aberta e integrada onde participem não só linguistas, mas que inclua também *“designers, developers and users of content encoding practices, and also many members of the society. It is already proved by a number of projects that lexicon building and maintenance can be achieved in a cooperative way”*.

Por outro lado, o estabelecimento de espaços multilingues de acesso à informação, como as redes sociais e as redes colaborativas, que se caracterizam pela sua heterogeneidade e multiculturalidade, cria a necessidade de existência de novos tipos de sistemas de organização do conhecimento que facilitem e tornem mais eficiente o acesso aos recursos de informação existentes e a partilha de conhecimento. O processo de desenvolvimento tecnológico, por um lado, e o aumento da cooperação a nível nacional e internacional, por outro, levam a que recursos de informação estruturados possam e devam ser reutilizados, integrados e disponibilizados de modo a simplificar o acesso à informação em diferentes línguas.

O uso e o acesso a ferramentas e recursos multilingues devem considerar, assim, para além da dimensão técnica, as dimensões organizacionais, económicas, legais, políticas e outras, devendo ser abordada a partir de uma variedade de perspectivas e envolver

diferentes comunidades, de modo a atender aos aspectos essenciais do multilinguismo no contexto do acesso e do uso de conteúdos digitais (Braschler *et al.*, 2010: 01).

Entre os recursos descritos, encontram-se os recursos terminológicos multilingues, que têm vindo a ser desenvolvidos a um ritmo acelerado ao longo das últimas décadas e que, mais recentemente, passaram a ser integrados em sistemas de tradução automática, memórias de tradução, sistemas de linguagens controladas, aplicações de linguística de *corpus*, entre outros.

Com a multiplicação das necessidades e propósitos, aqueles recursos passaram a ser utilizados de um modo ainda mais diversificado, sendo o seu uso cada vez mais recorrente também em sistemas de organização do conhecimento ou na construção de ontologias. O seu desenvolvimento tornou-se, por outro lado, uma actividade que passou a envolver, para além de terminólogos, investigadores de diferentes áreas que procuram aceder, organizar e disseminar informação e conhecimento de carácter específico.

A necessidade da existência de produtos terminológicos multilingues que apoiem tanto o aumento sustentado do comércio e do negócio electrónico como a transmissão de conhecimento entre comunidades, tem, por seu lado, contribuído para uma revisão e desenvolvimento das metodologias e práticas terminológicas⁴.

Para Cabré (2005: 04), entre os aspectos mais relevantes e que influenciaram a mudança da concepção teórica e de prática da terminologia encontra-se, precisamente, a multiplicação dos actos comunicativos plurilingues com recurso a línguas de diferentes estatutos sociopolíticos, entre países culturalmente diferentes e diferentes graus de desenvolvimento económico e tecnológico, a que acrescenta a diversificação das situações de comunicação especializada advinda quer das necessidades diversas de transmissão do conhecimento especializado, quer da multiplicação de temas que são alvo de intercâmbio internacional em resultado do crescimento acelerado do conhecimento.

A existência de terminologias de carácter multilingue, reutilizáveis e partilháveis, é, assim, indispensável à existência de uma comunicação especializada rigorosa e eficaz no seio

⁴ Aussenac-Gilles (2000), Cabré (2005), Wright *et al.* (2001), Gouadec (2003), Budin (2005), Calzolari (2005), Lenci (2008), entre outros, têm-se dedicado à análise desta problemática, concentrando a sua atenção nas propostas de abordagens, metodologias e ferramentas de base terminológica que contribuam para o acesso e a gestão mais eficiente do conhecimento em ambientes multilingues.

da sociedade da informação, dado o papel que representam enquanto meios de acesso e de disseminação do conhecimento. O trabalho e as metodologias terminológicas têm vindo, por isso, a evoluir consideravelmente, impulsionados pelos avanços técnicos e pelo interesse suscitado pela terminologia junto de áreas como a Inteligência Artificial (IA), entre outras⁵, desempenhando os terminólogos um papel de relevo em actividades profissionais onde as competências discursivas e a transferência e a organização do conhecimento se tornam necessárias.

1.2. Organização e partilha do conhecimento

As actividades de transferência, partilha e reutilização da informação e do conhecimento, bem como as abordagens multi e transdisciplinares adoptadas para a análise de cada domínio, levam à crescente necessidade de organização e representação do conhecimento em domínios de especialidade, os quais nem sempre são claramente delimitados, dada a multiplicação dos diferentes domínios científicos e técnicos e a sua interpenetração.

Esta dinâmica e a evolução dos diferentes domínios de conhecimento resultam da interacção constante entre indivíduos que prosseguem objectivos comuns no seio das organizações de que participam. Essa interacção, sobretudo quando estamos perante organizações multinacionais, tende a realizar-se, cada vez mais, em ambientes colaborativos e cooperativos disponíveis em linha, i.e., em espaços virtuais mediados pela internet, que podem ser vistos como espaços de comunicação computacional, ainda que humanizados.

A implementação deste tipo de processos colaborativos e a emergência do paradigma das empresas e das organizações virtuais situa-se na sequência natural dos processos de reestruturação dos paradigmas tradicionais da indústria resultantes dos avanços nas tecnologias da informação e comunicação. Por outro lado, é amplamente reconhecido que, no contexto económico actual, a cooperação entre empresas é um factor

⁵ Neste contexto, a extracção, estruturação e disseminação dos recursos terminológicos (Bourigault *et al.*, 2001), (Hamon & Nazarenko, 2002), (Bourigault & Aussenac-Gilles, 2003), (L'Homme, 2004), constituem uma contribuição de relevo para um grande número de abordagens em diferentes disciplinas, como a classificação de documentos ou a gestão de informação, entre outras, abordagens que assumem, cada vez mais, um carácter interdisciplinar e multilingue.

crucial de competitividade, em particular para as micro, pequenas e médias empresas (PME). Finalmente, as tendências do ambiente socioeconómico apontam para uma crescente complexidade na criação de valor tornando incontornável a cooperação entre as PME.

Estes ambientes colaborativos podem assumir várias formas, de que destacamos as redes colaborativas, que serão alvo da nossa atenção ao longo desta dissertação. Camarinha-Matos *et al.* (2005: 30) definem uma rede colaborativa⁶ como uma rede composta por uma variedade de entidades, e.g. organizações e pessoas, que possuem grande autonomia, se encontram geograficamente distribuídas e são heterogéneas em termos do seu

operating environment, culture, social capital and goals, where participants collaborate to (better) achieve common or compatible goals, and whose interactions are supported by computer network.

Segundo os autores, nas redes colaborativas existe uma intenção de colaborar que deriva da crença partilhada que, juntos, os membros da rede podem atingir objectivos que seriam inalcançáveis agindo de forma isolada ou que acarretariam custos elevados, incomportáveis para as organizações. Neste tipo de comunidades, o acesso e a disseminação do conhecimento surge ligado não tanto ao indivíduo em particular, mas ao grupo. Como descreve Cornejo (2003: 01) estas comunidades assumem a forma de grupos onde o conhecimento⁷ e as melhores práticas de um dado domínio são desenvolvidos, trabalhados e transmitidos através da interacção social dos seus membros, correspondendo a espaços onde o conhecimento tácito dos membros é transmitido a outros membros e onde este conhecimento é, normalmente, codificado e explicitado.

Nestes ambientes, a língua, meio por excelência da comunicação humana, e a terminologia, enquanto elemento nuclear à recolha, estruturação e disseminação do conhecimento, representam um papel de mediação cada vez mais importante na comunicação entre os vários interlocutores e na comunicação homem-máquina e um

⁶ Para Camarinha-Matos *et al.* (2005: 30) as redes colaborativas podem assumir várias formas, que incluem “*virtual organizations, virtual enterprises, dynamic supply chains, professional associations, industry clusters, professional virtual communities, collaborative virtual laboratories, and so forth*”.

⁷ Cornejo (2003: 02) distingue entre conhecimento, informação e dados. Para o autor conhecimento é “*the information required to satisfy a need (...) knowledge is such when it enables a person to do something. (...) If it is not good for anything but the person can understand it, it will simply be called ‘information’. If the person can’t make sense of it, it will be merely considered ‘data’*”.

elemento essencial ao desenvolvimento de uma estratégia de gestão de conhecimento que atinja o sucesso.

A gestão do conhecimento tem como objecto de estudo a identificação, aquisição, organização, criação, uso e disseminação do conhecimento nas suas muitas formas. De modo a desenvolver e utilizar estratégias e tecnologias de apoio à gestão do conhecimento, temos necessidade de aprofundar a nossa compreensão quanto ao modo como os indivíduos, grupos e organizações verbalizam, usam e disseminam o conhecimento.

De facto, a representação sistemática e dinâmica da informação e do conhecimento interessa a todos os intervenientes de uma comunidade, dada a obrigatoriedade da partilha de conhecimento com coerência entre os diferentes agentes e a cada vez maior relevância e urgência do uso de dados estruturados e normalizados em sistemas de informação e comunicação, que permita transformar a informação existente numa organização em conhecimento acessível, assimilável, disseminável e reutilizável.

Como apontam Davenport e Prusak (2000), para que uma estratégia de gestão de conhecimento surta efeito, torna-se necessária a existência de uma cultura de transferência de conhecimento adequada, que permita ultrapassar os problemas causados pelas diferentes culturas, terminologias e modelos de referência, através da criação de uma terminologia comum e de uma cultura organizacional partilhada entre aqueles que querem cooperar dentro de um esquema de gestão do conhecimento. Tal como afirmam Budin e Melby (2000: 02),

a common and clear language with shared meanings is not only the basis of any particular culture, but also the prerequisite for any successful terminology sharing strategy within and among various organizations.

A construção e a transmissão do conhecimento ocorrem através de uma actividade lógica e discursiva, desenvolvida pelo recurso a signos (Rey, 1995: 47). Para o autor, na nossa experiência cultural e na construção do conhecimento recorreremos aos signos linguísticos, e, para este propósito, especialmente aos nomes. Nomeamos para diferenciar, reconhecer e, finalmente, para saber.

A língua é, como descreve Rogers (2008: 113), o primeiro meio para um ser humano adquirir conhecimento sobre o mundo, transmitir representações mentais e torná-las

públicas e acessíveis intersubjectivamente e representa o instrumento nuclear à existência de um repositório colectivo de conhecimento (*collective knowledge reservoir*) a ser transmitido de geração em geração, actuando também como um meio de categorizar as experiências culturais, o pensamento e os comportamentos para os seus falantes, pelo que, como concluí a autora, a língua e a cultura estão, de facto, intimamente interrelacionadas ao nível semântico, onde o vocabulário de uma língua reflecte a cultura partilhada pelos seus falantes.

O conhecimento⁸ não pode, assim, ser separado do contexto, da experiência, da cultura, da língua e do potencial criativo desta, sob pena de não ser correctamente apreendido, organizado e disseminado. Budin (2007: 70) enfatiza essa ligação ao afirmar que:

language change is a function of conceptual change and in turn inspires further conceptual change.

A língua e a terminologia assumem, assim, um papel fundamental na descoberta, criação e transmissão do conhecimento, bem como na sua eficiente conceptualização, representação e reutilização. Como afirma Temmerman (2000: 62), ao destacar a importância da língua na compreensão do mundo e no processo de comunicação, a compreensão da língua é inseparável da compreensão do mundo.

1.2.1. Conceito

A importância do conceito está patente nos muitos estudos teóricos e definições de que foi alvo nas últimas décadas, o que nos coloca perante a evidência das inúmeras abordagens possíveis à sua significação, como retracta Jackendoff (1995: 21):

⁸ O conhecimento é definido por Hessen (1987: 27) como uma *determinação do sujeito pelo objecto*, em que o determinado não é o sujeito, pura e simplesmente, mas a sua representação mental do objecto. Esta imagem, constituída através de um processo de abstracção e de generalização, é o meio que permite ao sujeito cognoscente apreender o objecto. Como explicita Remígio (2010: 20), este processo implica um dualismo entre o sujeito que apreende e o objecto que é apreendido, em que, porém, e tal como refere Hessen (1987:27), “o objecto não é arrastado (...) para dentro da esfera do sujeito, mas permanece, transcendente a ele”, sendo a imagem do objecto, i.e., a sua representação, que é efectivamente iminente ao sujeito. Consequentemente, sendo a realidade e o conhecimento sobre essa realidade distintos, a representação do objecto é passível de ser alterada. Contudo, o conhecimento não corresponde apenas a uma actividade cognitiva do Homem. É também o produto dessa actividade, fruto de uma construção mental do sujeito, o qual está inserido num contexto social, histórico e temporal específico. (Remígio, 2010: 21)

Asking a psychologist, a philosopher, or a linguist what a concept is is much like asking a physicist what mass is. An answer cannot be given in isolation. Rather, the term plays a certain role in a larger worldview that includes the nature of language, of meaning, and of mind. (...) the "correctness" of a particular notion of concept cannot be evaluated without at the same time evaluating the world view in which it plays a role.

Para representar o conhecimento precisamos de identificar os conceitos, sendo de primordial importância compreender como proceder à sua representação. No que respeita à terminologia, a importância e centralidade do conceito para a estruturação do conhecimento dos indivíduos foram reconhecidas e descritas por Wüster, na Teoria Geral da Terminologia, que classificou o *conceito* como o ponto de partida, a unidade de conhecimento fundamental em todo o processo terminológico. Para Wüster (1985: 01) o trabalho terminológico parte do conceito: *geht von den Begriffen aus. Sie zielt auf scharfe Abgrenzung zwischen den Begriffen*. Esta centralidade do conceito é destacada por Picht & Draskau (1985:36) que admitem a existência de um consenso quanto à centralidade ocupada pelo conceito na teoria da terminologia.

No entanto, tal consenso tem sido alvo de uma ampla discussão teórica, onde se analisa a definição de conceito no seio dos estudos terminológicos à luz das diferentes metodologias propostas. Esta discussão tem sido alvo de contributos variados e distintos que envolvem um grande número de autores como Wüster (1985); Felber (1987); Picht & Draskau (1985); Sager (1990); Hohnhold (1990); Temmerman (2000); Roche (2007, 2008, 2012); entre outros.

O desafio que se coloca à definição de *conceito* deriva, como descreve Antia (2000: 81), do carácter eclético da terminologia enquanto disciplina, i.e., do facto de ser uma ciência interdisciplinar e incorporar conhecimento de outras disciplinas, algumas das quais com um longo historial de análise e discussão sobre o conceito, como a filosofia, a linguística, a psicologia ou a engenharia do conhecimento, onde, como descreve, existe ou existiu uma preocupação com os conceitos no contexto de uma ou na combinação das seguintes acções ou realidades: *"perceiving it, (re)creating it, stabilising it, organising it, and communicating it"*.

A terminologia é, por outro lado, como descreve Antia (2000: 81), um “*phenomenon of specialised subject areas*”, o que implica, na sua perspectiva, que as diferentes descrições do conceito devem reconhecer que os domínios de especialidade:

work with different ontologies and are underpinned by different epistemological positions — a diversity that can be expected to affect how concept creation and description is viewed.

Não podemos ignorar, também, que esta demanda na tentativa de encontrar uma definição mais consensual recebeu um contributo significativo resultante da forte evolução em termos de metodologias e práticas terminológicas, sobretudo nas últimas duas décadas, reflexo das necessidades advindas das alterações no pensamento da ciência em geral e da necessidade de enfrentar o que alguns autores percebem como as restrições que estariam presentes nas premissas originais da terminologia (Temmerman 2000)⁹.

Esta diversidade requer, por parte da terminologia, uma abordagem à teoria do conceito sem pretensões de universalidade, dependendo a definição e descrição de um conceito de um suporte teórico variável e dos objectivos a que se propõem (Santos, 2010: 85).

Tendo em conta a diversidade de contributos, começamos a nossa análise com a definição proposta por Wüster (1985: 07), para quem o conceito é um *elemento do pensamento (Denkelement)*, definido pela posição que ocupa num sistema de conceitos. Para Wüster, os conceitos são representações mentais que ajudam a estruturar os objectos do mundo real, sendo na sistematização desses conceitos¹⁰, que reside a essência do trabalho terminológico.

⁹ Temmerman (2000: 01) sumariza essas premissas do seguinte modo: “*that concepts are clear-cut and can be defined on the basis of necessary and sufficient conditions, that univocity of terms is essential for unambiguous and therefore effective and efficient communication, and that figurative language and change of meaning are linguistic subjects which are of no concern to Terminology as Terminology restricts itself to the onomasiological perspective*”, i.e., que os conceitos têm contornos precisos e um lugar definido num sistema conceptual, que a terminologia parte do estudo dos conceitos, seguindo-se o estudo dos termos (abordagem onomasiológica) e que o conceito e o termo se relacionam de modo biunívoco, sendo essencial que a relação entre conceitos e termos seja acordada e normalizada de modo a diminuir a ambiguidade e a contribuir para uma comunicação eficiente.

¹⁰ Tal como observa Costa (2010), apesar de colocar a sistematização dos conceitos ao nível conceptual, para Wüster (1985: 78) esta sistematização pode também situar-se ao nível denominativo: “*Jede Sprache, und erst jede Terminologie, bildet ein Begriffs- und Benennungssystem. Genau genommen bilden nur die Begriffe immer ein System, auf Grund der zwischen ihnen bestehenden Beziehungen. Die Benennungen und anderen Zeichen*

Felber (2001: 43) aprofunda a perspectiva de Wüster e define o conceito não só como uma representação mental (uma abstracção) de um objecto, mas também como o instrumento que possibilita essa apreensão: *“der Begriff ist Mittel und Werkzeug zur Erfassung und Darstellung von Gegenständen”*. Para Felber o conceito é um *objecto de pensamento* (Denkgegenstand)¹¹. No entanto, para Felber o conhecimento não se reporta a objectos, propriedades ou relações individualmente, mas a factos, representados mentalmente através de juízos e expressos por proposições, sendo o conceito perspectivado como um elemento constituinte da unidade do conhecimento¹², que é o juízo (Remígio, 2010: 128).

Laurén *et al.* (1998: 133) defendem que, não obstante a problemática que rodeia a definição de conceito, ora entendido como uma *unidade de pensamento*, ora como uma *unidade de conhecimento*, estas abordagens não se opõem, representando antes funções do conceito, funções a que acrescentam uma outra, a do conceito como unidade de cognição.

Os autores descrevem o conceito como uma unidade complexa, ligada ao pensamento, ao conhecimento e ao processo de cognição humana¹³. Para Laurén *et al.* (1998), pensamento, conhecimento e cognição estão em grande proximidade, interligam-se e condicionam-se, mas não são idênticos¹⁴. Destacam, nesta interligação, a unidade de conhecimento *“die Intension einer Wissenseinheit [ist] größer (...) als die einer Denkeinheit”* (Laurén *et al.*, 1998: 136).

Para os autores, enquanto *unidade de conhecimento*, o conceito é composto pela totalidade das suas características, tal como concebidas por uma comunidade de

dagegen bilden im allgemeinen ein System nur mittelbar, d.h. dadurch, dass sie den Gliedern eines Begriffssystems zugeordnet sind.”

¹¹ *“Der Begriff ist ein elementares Denkgebilde, das einem Gegenstand zugeordnet ist und diesen im Denken vertritt”*. (Felber, 2001: 45)

¹² Este tipo de conhecimento enquadra-se no conhecimento de tipo *proposicional* (*propositional knowledge*), tal como é descrito por Galinski e Picht (2001: 43), que distinguem entre aquele e o *conhecimento conceptual* (*conceptual knowledge*). O conhecimento conceptual *“involves knowledge pertaining to a single concept in the terminological sense, and the latter relates to all forms of knowledge that are based on propositions, in which at least two concepts are placed in relation to one another so that they make a meaningful statement”*.

¹³ Como refere Remígio (2010: 130), para os autores, da cognição enquanto processo – *Erkenntnisprozeß* - poderá resultar uma nova configuração das características que constituem o conceito, sendo o conceito, nesta linha de conta, perspectivado como objecto sujeito a modificações resultantes do processo cognitivo.

¹⁴ *“Denken, Wissen und Erkenntnis Grössen sind, die in einem engen Verhältnis zueinander stehen, die einander bedingen, aber nicht identisch sind”* (Laurén *et al.*, 1998: 133).

especialistas, num determinado contexto, sendo que o conjunto dos conceitos inter-relacionados representa o conhecimento dos especialistas acerca de uma determinada realidade.

Na norma ISO/FDIS 704 (2009: 02), que estabelece os princípios e métodos do trabalho terminológico, o conceito é considerado como uma representação mental de um objecto pertencente a um domínio ou contexto especializado. De acordo com a norma, o conceito pode ser entendido não só como uma unidade de pensamento (*unit of thought*), mas também como uma unidade de conhecimento (*unit of knowledge*).

Um conceito resulta, segundo a norma, de uma combinação única de características¹⁵ abstraídas a partir das propriedades de um objecto¹⁶ ou de um conjunto de objectos. Os objectos percebidos como partilhando as mesmas propriedades são agrupados em unidades.

Once similar objects, or occasionally a single object, are viewed as a meaningful unit of knowledge within a branch of human knowledge, the properties of an object, or those common to a set of objects, are abstracted as characteristics that are combined as a set in the formation of a concept. (ISO 704:2009(E), 2009: 04)

A norma estabelece uma distinção entre dois tipos de conceitos: o conceito individual, que corresponde a um objecto único, e o conceito geral que corresponde a um conjunto de dois ou mais objectos que formam um grupo em função das suas propriedades comuns.

Esta distinção, já anteriormente estabelecida pela norma ISO 704:2000, é para Roche (2012: 20), de evitar, uma vez que, na sua perspectiva, o conceito é uma *unidade de conhecimento* que envolve uma pluralidade de coisas, qualquer que seja o número de objectos que representa. Para o autor,

The definition of a concept does not depend on the cardinality of its extension. Furthermore, considering any unique object as an individual concept is not only unnecessary in terminology work but it raises several issues.

¹⁵ As características são descritas como qualificadoras, restringindo o significado de um conceito superordenado.

¹⁶ Abstracção é definida como sendo o processo de reconhecimento de um determinado conjunto de características comuns *“in an individual set of objects and, on that basis, forming a concept of that set of objects”* (ISO/FDIS 704:2009(E), 2009: 04).

Roche (2008: 55) descreve o conceito como um conhecimento “*portant sur une pluralité de choses, une « représentation intellectuelle permettant de viser le réel suivant des déterminations abstraites et générales et non dans sa singularité concrète »*. Para o autor, compreender a variedade de objectos que compõem a realidade torna-se possível se entendermos o conceito enquanto *unidade de compreensão*.

Na perspectiva do autor (2008: 55), mais do que uma *unidade do pensamento* (ISO 704: 2000), um conceito é antes uma *unidade de compreensão (unit of understanding)*, entendida como unidade que permite apreender a realidade na sua diversidade. Roche distingue, no entanto, a sua definição da proposta por Temmerman (2000: 34)¹⁷, que descreve como confundindo classificação e conceptualização.

Para Roche (2008: 70; 2012: 29) a noção de conceito, enquanto *unidade de compreensão*, permite a apreensão da diversidade dos objectos que povoam a realidade, desempenhando, para este fim, uma dupla função: a de compreender o que é uma coisa e a de a descrever, sendo que a primeira define a coisa em termos das suas características essenciais¹⁸ e a segunda descreve-a como um elemento único de conhecimento sob a forma de atributos (*valuated attributes*).

A diversidade de perspectivas e de abordagens, que continua a construir-se, está já presente na visão de Antia (2000: 82), que apresenta uma definição de conceito baseada numa abordagem que descreve como policêntrica e flexível. Para o autor, que procura seguir um ponto de vista unificador em relação às diferentes definições de conceito, o conceito é uma unidade do *pensamento, conhecimento ou cognição* que:

- i) *though dependent on language is independent of a given language; it can however be influenced by a variety of socio-cultural factors that correspond to linguistic boundaries;*
- ii) *is a mental representation, reduction or (re)interpretation of reality that is perceptible, imperceptible or that was previously non-existent;*

¹⁷ Partindo de uma perspectiva sociocognitiva, Temmerman (2000: 34) aponta a necessidade de utilizar meios mais complexos para a identificação dos conceitos e propõe a substituição da abordagem conceptual por uma abordagem de compreensão (*We replace the meaning approach by an understanding approach*), em que a *unidade de compreensão* assumiria o papel de unidade central em Terminologia.

¹⁸ Para Roche (2012: 29), as características essenciais derivam da razão, participam na definição dos conceitos e organizam-nos num sistema: “*a skeleton – on which are hung the attributes describing the objects as they are perceived and whose values express contingent knowledge*”.

- iii) *is comprised of characteristics that are (at some point, at least) negotiated within a specialized knowledge community;*
- iv) *typically enters into some (organic or logical) relationship with other concepts;*
- v) *can exist without symbols (whether linguistic or non-linguistic), but requires symbols for purposes of communication.*

Como observa Antia, esta visão policêntrica procura ser uma súpula das diferentes definições presentes na literatura. É, no entanto, uma visão que importa reter, dada a sua abrangência, o modo multifacetado como define e descreve o conceito, como afirma a sua independência em relação à língua, ainda que reconhecendo a influência desta, e, finalmente, como inclui a negociação do seu significado no seio de uma comunidade, numa posição próxima da de Laurén *et al.* (1998).

Tal como postula a norma (ISO/FDIS 704:2009(E), 2009: 08), os conceitos não existem enquanto unidades de conhecimento isoladas, mas apenas nas relações que estabelecem com outros conceitos, sendo estas relações passíveis de serem constantemente criadas e refinadas: *“Our thought processes constantly create and refine the relations between concepts, whether these relations are formally acknowledged or not”*.

Tradicionalmente, o terminólogo ordena os termos de acordo com uma estrutura conceptual, com o objectivo de delimitar um domínio específico. Para tal, identifica e estabelece relações entre os conceitos pertencentes ao domínio em análise. A estruturação de um conjunto de conceitos a partir das relações que estabelecem entre si dá origem a um sistema de conceitos.

Na norma é feita a distinção entre dois tipos de relações entre conceitos: as relações hierárquicas, subdivididas em relações genéricas e relações partitivas, e as relações associativas. Como descreve Roche (2012: 23), o primeiro tipo de relações desempenha um papel central, dado que estruturam o sistema conceptual, permitindo-nos, assim, a sua compreensão e o domínio sobre a sua complexidade, enquanto o segundo tipo expressa uma conexão considerada não-hierárquica entre conceitos, que surgem por via da experiência, relações tais como as que envolvem uma causa e o seu efeito ou produtor e produto, entre outras.

O processo de organização de conceitos num sistema conceptual obriga a que se tenha presente o domínio a que pertence o conceito e a que se considerem as expectativas e objectivos dos utilizadores finais: *“The subject field shall act as the framework within which*

the concept field, the set of thematically related but unstructured concepts, is established.” (ISO 704, 2009: 08). Como descreve Roche (2008: 56),

Deux objets relèvent d’un même concept, non pas parce qu’ils partagent les mêmes caractéristiques, mais inversement, c’est parce que les caractéristiques essentielles du concept leur sont applicables. Ce sont ces caractères essentiels qui font le sujet « autre » et qui, s’ils lui étaient retranchés, il ne serait plus ce qu’il est. En cela, ils ne peuvent être soumis « au plus ou au moins ». Les caractères essentiels définissent les concepts – au sens où ils permettent de comprendre la chose – et les organisent en système.

Torna-se assim importante reconhecer os conceitos que compõem um determinado domínio específico e, com base nas características identificadas, proceder à sua estruturação e à representação do conhecimento do domínio.

Segundo Depecker (2002: 87-88) o conceito é multidimensional. Na sua perspectiva, se admitirmos o facto de que o conceito pode ser abordado de pontos de vista diferentes, e portanto, a partir de critérios diferentes que permitem determinar a sua dimensão, devemos considerar que um conceito é:

un élément à plusieurs dimensions et que la dimension choisie varie à chaque fois en fonction de l’angle de vue retenu dans le travail terminologique. (Depecker, 2002: 86)

Esta multidimensionalidade do conceito pode reflectir-se nas suas definições e nas denominações que lhes correspondem (Depecker, 2002: 121). Também Budin (2007: 70)¹⁹ que tem em conta os pontos de vista de Beaugrande (1998, 1991), Toft (1990) e Temmerman (2000), confere ao conceito um carácter multidimensional. Segundo o autor, os termos não são, necessariamente, fixos no seu significado, uma vez que, dependendo da situação em que são usados, o seu conteúdo conceptual pode variar.

Segundo Kageura (1997: 120) esta multidimensionalidade do conceito leva a que também o sistema conceptual deva ser entendido como multidimensional. Na organização do sistema conceptual, o autor (1997: 128) estabelece a distinção entre dois tipos de multidimensionalidade:

¹⁹ Esta perspectiva está próxima da de Michalski (1993: 146), para quem os conceitos não representam, “*fixed, well defined structures, but as flexible, context modifiable, are background-knowledge-dependent units of our knowledge*”.

- *multidimensionality introduced by the co-existence of different means of concept classification, i.e., generic/specific, part/whole, and type/value,*
- *multidimensionality introduced by the application of different types of characteristics or facets at the same level in the generic/specific structure.*

Esta multidimensionalidade permite, assim, que se combinem diferentes tipos de relações conceituais no mesmo sistema de conceitos.

Importa, finalmente, na análise do que é um conceito, distinguir entre conceito e significado, muitas vezes tidos como sinónimos. Depecker (2002: 30) separa o plano conceptual, do plano linguístico, concretamente através da distinção entre conceito e significado. Para o autor o conceito não se resume ao significado: um e o outro distinguem-se, mesmo quando possam vir a ser confundidos na língua.

Apesar da distinção que estabelece, Depecker (2002: 56) afirma que entre conceito e significado não existe uma separação claramente demarcada, por um lado, porque existe uma interacção evidente e permanente entre os dois, o que nos leva a confundi-los e, por outro, porque é necessário admitir que existe *“du flou aussi bien dans la pensée que dans la langue, ce que tout scientifique admettra”*, pelo que as delimitações que estabelecemos são, frequentemente, pouco satisfatórias.

No entanto, esta distinção entre conceito e significado tem como vantagem, segundo o autor, permitir ao signo linguístico preservar toda a sua riqueza, particularmente se tivermos em conta uma perspectiva terminológica, permitindo também que se observem as ligações entre conceito e signo. Tendo em conta este ponto de vista, Depecker aproxima-se da vertente conceptual, dado que perspectiva o conceito como independente da língua, como um elemento extralinguístico.

Prandi (2011) descreve o léxico de uma língua, como não sendo mais do que *um contentor de conceitos*, sendo que o léxico constitui, sobretudo, uma estrutura complexa e estratificada de modo a dar uma forma aos conceitos, os quais não são mais do que significados. Mas, para o autor, no momento em que um conceito se torna o significado de um signo, o conceito, mesmo o mais imediatamente acessível à experiência directa, *“s’engage dans une dérive aux issues imprévisibles qui peut le laisser presque intact aussi bien que le restructurer d’une façon plus au moins radicale”*.

Para Fernández (2007: 43), o reconhecimento da distinção entre conceito e significado permite:

- *integrar las tendencias opuestas que se dan en el lenguaje especializado, a saber, la necesidad de estabilidad y una cierta universalidad de las nociones y el respeto hacia las particularidades derivadas de la multiplicidad de percepciones asociadas a lenguas y culturas.*
- *justificar y explicar el fenómeno de la variación denominativa entendida como la convivencia de varios signos lingüísticos que hacen referencia a un mismo concepto.*

Como defende Depecker (2002: 115), tendo em conta esta diferença entre significados nas diferentes línguas, o significado de um termo numa língua nem sempre corresponde, a não ser de modo imperfeito, ao significado de um termo numa outra língua, pelo que qualquer análise terminológica deve, necessariamente, englobar uma análise de nível conceptual, de modo a melhor preencher as eventuais lacunas conceptuais que se detectem, perspectiva com a qual concordamos.

1.3. Conceptualização

O uso de um conceito segue um ciclo de vida. Como descreve Santos (2010: 84) com base em Antia (2000: 83), um conceito é relevante enquanto unidade de pensamento quando inicia o seu ciclo de vida. No momento em que está totalmente formado e aceite pela comunidade de especialistas, transforma-se em unidade de conhecimento, dando-se a passagem a unidade lógica com recurso a uma linguagem de representação formalizada para fins de manipulação computacional.

A conceptualização do domínio constroi-se a partir do conceito. Como descreve Costa (2009: 02), o especialista apreende o conceito delimitando-o:

em função da sua própria formação, em função da visão que tem do mundo. Ele olha e reflecte sobre o conceito, sob um determinado prisma e tentará retrata-lo o melhor possível no momento de transmitir o conhecimento que tem acerca desse conceito²⁰.

²⁰ Este processo de apreensão de um domínio (i.e. dos objectos que povoam uma “realidade”) e a sua conceptualização remetem-nos, segundo Roche (2008), para a categoria dos pensamentos “universais” próprios a uma cultura.

A conceptualização resulta, então, de um processo que leva à extracção e generalização de informação relevante baseada na experiência do indivíduo e pode ser definida como o conhecimento informal relevante que podemos extrair e generalizar a partir da experiência, da observação ou da introspecção (Prévot, 2010: 842). Situa-se, nas palavras de Barros (2007: 12), entre a apreensão da realidade e a sua interpretação.

Quando analisámos a conceptualização no âmbito dos sistemas de informação e a enquadrámos no processo de construção de uma ontologia, deparamo-nos, naturalmente, com a definição de Gruber (1993a) para quem uma ontologia é uma *especificação explícita de uma conceptualização*. Esta definição baseia-se na ideia de conceptualização vista como uma visão do mundo que pretendemos representar, correspondendo a conceptualização ao processo pelo qual a mente humana forma o seu conhecimento sobre uma parte da realidade. Este conhecimento deve corresponder a uma representação mental livre de propriedades acidentais e baseado nas características essenciais dos elementos.

Gruber (1993a, 1995) desenvolve a noção de conceptualização a partir de Genesereth e Nilsson (1987), para quem o conhecimento representado formalmente parte de uma conceptualização, que descrevem como sendo o conjunto dos objectos, conceitos e outras entidades:

that are assumed to exist in some area of interest and the relationships that hold among them. A conceptualization is an abstract, simplified view of the world that we wish to represent for some purpose.

A conceptualização resulta, assim, de uma análise e estruturação do conhecimento de um domínio particular, elaborada de acordo com um propósito específico, sendo composta pelo conjunto dos conceitos identificados e pelas relações conceptuais que estes estabelecem entre si.

No âmbito das ciências da computação o propósito principal de uma conceptualização é o de criar ontologias tendo por base um modelo concreto do mundo e, através de um processo descritivo, criar um modelo abstracto para capturar a sua essência. A conceptualização cria, no fundo, as condições para a transferência das representações mentais para a ontologia (Sánchez *et al.*, 2005: 07).

Para os autores, a conceptualização representa uma versão de um domínio que procura representar a essência dos objectos, i.e. descrever as características mínimas essenciais que fazem um objecto ser ele próprio e diferente de todos os outros. Inclui, por isso, um processo de abstracção, a partir do qual os criadores da ontologia observam o mundo (ou o domínio), analisam as suas observações e usam o seu conhecimento anterior, produzindo um resultado, que define um conceito (Sánchez *et al.*, 2005: 07).

Daqui podemos concluir que o processo de conceptualização corresponde a um tipo específico de abstracção que tem como objectivo identificar um conjunto de conceitos de um domínio, correspondendo a conceptualização obtida neste processo a um subconjunto do conjunto de todas as construções mentais que um indivíduo pode construir.

A definição de conceptualização proposta por Sánchez *et al.* (2005) deve, a nosso ver, ser acrescida do *propósito* que preside ao processo de abstracção, tal como defende Gruber (1993b: 01). Para o autor, uma conceptualização corresponde a uma visão abstracta e simplificada do mundo que é nosso intuito representar de acordo com um propósito específico, sendo que todos os sistemas ou agentes baseados em conhecimento estão, explícita ou implicitamente, comprometidos com algum tipo de conceptualização de modo a constituir um corpo de conhecimento representado formalmente.

Sánchez *et al.* (2005: 08) consideram ainda que, se as representações de conceitos correspondem a um tipo de representação mental, então as ontologias são um conjunto de modelos cujo objectivo é o de representar conceitos e todas as interpretações possíveis de acordo com o domínio a que os conceitos pertençam.

Guarino (2009: 02), por seu lado, alerta para o facto de o modo como uma ontologia específica uma conceptualização não ser um processo único ou simples, estando dependente de várias condicionantes, entre as quais “(1) *the richness of the domain of discourse*; (2) *the richness of the vocabulary chosen*; (3) *the axiomatization*”.

Importa, aqui, sublinhar, que não se deve confundir representação do conhecimento com conhecimento, como alerta Santos (2010: 94). Na opinião desta, dada a impossibilidade de existir “*uma conceptualização ‘correcta’ ou ‘verdadeira’ sob determinada realidade, reflecte-se uma visão do mundo definida por um determinado indivíduo ou grupo de indivíduos, que é válida para um objectivo específico*”. Para Santos (2010: 86-87), o processo

de conceptualização assume uma função activa na definição do significado dos conceitos, uma vez que, como afirma, os conceitos variam em função *“das conceptualizações e do ponto de vista de quem o descreve e/ou define, representando os termos que os designam um meio de adquirir, criar, comunicar, armazenar, representar e operacionalizar conhecimento”*.

1.3.1. Especificação semiformal da conceptualização

Embora exista um número crescente de ferramentas e recursos semânticos disponíveis, surgem ainda vários problemas quando procuramos estabelecer uma conceptualização comum de uma dada realidade, como reconhece Pereira (2010: 01), uma vez que:

(i) apesar da evolução das tecnologias semânticas, é virtualmente impossível criar artefactos semânticos que consigam responder completamente a todas as variações possíveis que podem ocorrer em diferentes contextos e situações de negócio (sendo estes cada vez mais dinâmicos); (ii) apesar de todos os esforços para definição de normas, sabe-se que existe um tipo de “resistência social” na aceitação de normas semanticamente orientadas.

Esta visão parte dos problemas e falhas apontados ao processo de construção de ontologias por autores como López *et al.* (1999) que vêem como necessário, por um lado, o reforço do desenvolvimento de metodologias que apoiem a fase complexa de aquisição de conhecimento e, por outro, o desenvolvimento de ferramentas que apoiem a conceptualização do conhecimento e que, tendo por base esta conceptualização, geram a ontologia.

Staab (2008) aponta um outro problema, que advém do facto de a noção de conceptualização partilhada não estar ainda bem explorada, nem bem entendida, nem ser compatível com a maioria das ferramentas de engenharia de ontologias, apesar dos diferentes graus de formalização estarem já bem investigados e integrados em diferentes tecnologias baseadas em ontologias.

de Moor (2005), por outro lado, chama a atenção para a necessidade da investigação em engenharia de ontologias se concentrar mais no desenvolvimento de processos de negociação de significado, processo que descreve como parte integrante do processo de construção colaborativa de uma conceptualização.

Este processo de negociação de significado é descrito por de Moor (2005) como o desenvolvimento de *“ontology-guided, community-oriented processes for agreeing upon and reaching the appropriate amount of consensus on terminology and concept definitions”*, processos que entende como essenciais ao desenvolvimento efectivo de serviços colaborativos no contexto de um comunidade em que coexistem diferentes significados e interesses. Por outro lado, para o autor, o problema de como chegar a um consenso sobre o conteúdo de uma ontologia num ambiente distribuído, tal como acontece no caso das redes colaborativas, antes descritas, não se encontra ainda resolvido.

Tendo em conta estas lacunas, Pereira (2010: 24,25) enquadra a conceptualização no contexto do espaço partilhado por uma comunidade e descreve-a como um processo dinâmico de construção de significado, definindo o processo de conceptualização como *“o conjunto de actividades que nos permitem obter uma conceptualização de um determinado domínio de aplicação, partilhada por um conjunto de actores”*.

Na definição que propõe, a autora inclui, no processo de conceptualização, o processo de partilha de conhecimento sobre um domínio, levado a cabo por um conjunto de actores que formam um comunidade, processo que entendemos como essencial para a obtenção de uma conceptualização efectiva e partilhada do conhecimento do domínio, sobretudo quando desta partilha resulta uma ontologia representada em mais do que uma língua, como procuraremos demonstrar ao longo desta dissertação.

O processo de conceptualização pode, segundo Pereira et al. (2012), ser de ordem individual ou colaborativa. No que se refere à conceptualização individual, descrevem o processo de conceptualização de uma dada realidade como sendo *“uma colecção de actividades cognitivas ordenadas, que tem como input informação e conhecimento interna e externamente acessível ao individuo e como output uma representação conceptual interna ou externa”*. O processo de conceptualização colaborativa, por seu lado, é descrito como uma *“conceptualização que envolve mais do que um indivíduo, produzindo uma representação conceptual obtida por consenso”*. Para além da representação conceptual

individual, a representação conceptual colaborativa envolve actividades sociais que incluem a negociação do significado e a gestão das actividades de desenvolvimento do próprio processo colaborativo (Pereira, 2010: 24,25).

Pereira & Soares (2008: 614) situam este processo de conceptualização na fase inicial da construção da ontologia, ou seja, numa fase que designam de *fase de especificação semiformal*, em que não se recorreu ainda a uma representação formal (através de linguagens como o *owl* ou o *rdf*).

Para Pereira (2010: 24,25), a conceptualização é uma actividade vital à partilha de informação e conhecimento no contexto de uma rede colaborativa, constituída por equipas provenientes de diferentes organizações, com a sua própria cultura, experiências e visão do mundo. Segundo Pereira & Soares (2008: 606), a conceptualização é a base para a organização do sistema de informação e conhecimento, sendo fundamental para estruturar, armazenar e extrair informação e essencial à compreensão efectiva das práticas colaborativas. Para os autores, a actividade de conceptualização permite a criação de um modelo conceptual que reflecte *“the semantic agreement established that will be the keystone for information and knowledge sharing”* (*idem, ibidem*).

O sucesso de uma ontologia depende, assim, directamente da conceptualização, que, tal como defende Pereira (2010: 157), é *um processo social*, i.e., que necessita de uma presença social elevada na criação de um modelo conceptual que reflecta os acordos semânticos²¹ que serão a base para a partilha de informação e conhecimento, tal como é descrito na figura seguinte.

²¹ Acordos obtidos entre especialistas, no desenrolar do processo de conceptualização, relativos ao significado dos termos de um domínio específico.

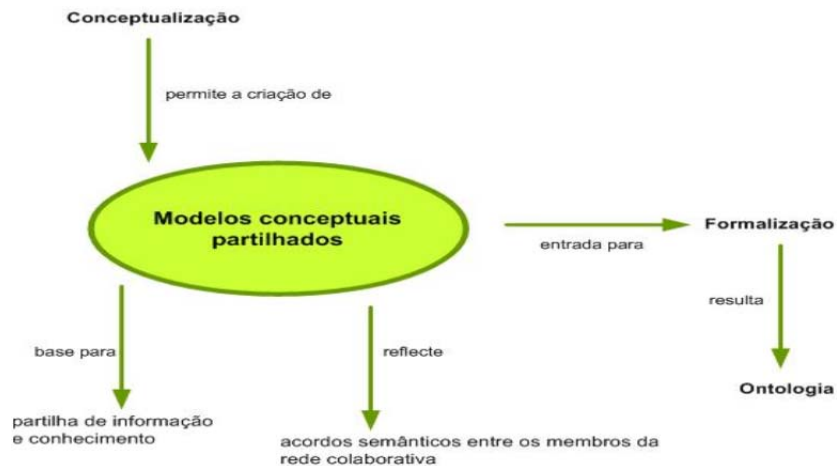


Fig. 1 - Relação entre actividade de conceptualização e desenvolvimento colaborativo de ontologias. (Pereira 2010: 30)

Para a autora, as actividades de conceptualização incluem a *identificação, análise e negociação* dos conceitos a incluir na ontologia, a análise das suas características, definição e relacionamentos, bem como a *“organização e estruturação do conhecimento usando representações externas independentes dos artefactos de implementação”* (Pereira, 2010: 03), sendo o resultado do processo de conceptualização constituído por uma representação semiformal do conhecimento, sob a forma de mapas conceptuais²².

A conceptualização ocorre, assim, na fase que antecede a especificação formal da ontologia, resultando este processo num modelo conceptual que reflecte os acordos semânticos estabelecidos para a partilha de informação e conhecimento (Pereira, 2010: 13) e implica a existência de uma estrutura conceptual e de um processo de representação das estruturas conceptuais, que conduza à externalização de uma estrutura conceptual através da sua representação.

Dos diversos formalismos de representação de conhecimento existentes, a representação e especificação da conceptualização pode ser realizada com recurso a mapas de conceitos (Pereira, 2010: 25,26). Estes mapas de conceitos, cuja utilização analisaremos com maior profundidade no ponto 3.6, correspondem a um formalismo gráfico de representação de conhecimento que permite a captura, construção, reflexão, refinamento, comunicação, colaboração e transferência de conhecimento.

²² Roche (2006a) descreve esta representação como um formalismo semi-formal: *“we can represent the denoted concepts and their organisation in a semi-formal formalism as a network of concepts”*.

Esta forma de modelação conceptual possibilita a descrição de modo semiformal de aspectos do mundo físico e social que nos rodeia com o intuito de permitir a compreensão e a comunicação. Desta actividade resulta o que Pereira (2011: 19) designa de *representações externas*, correspondentes a representações visuais do conhecimento, as quais são partilhadas por uma determinada comunidade, obtendo a sua coerência de acordo com o conhecimento científico dessa comunidade.

O facto desta externalização poder, nesta fase, ser expressa com recurso à língua natural, permite aos especialistas do domínio, por outro lado, afastarem-se dos paradigmas de modelação das ontologias e das linguagens de representação subjacentes e concentrarem-se nas (suas) necessidades de modelação do conhecimento do domínio.

Neste contexto, acreditamos que a hipótese dos membros de uma comunidade colaborativa especificarem uma ontologia com recurso à língua permite que a actividade de construção de uma ontologia possa ser percebida como uma actividade mais acessível. Acreditamos, ainda, que se os especialistas puderem expressar as suas necessidades de representação do conhecimento na sua própria língua de trabalho, a eventualidade de contribuírem para a construção e validação das ontologias pode tornar-se mais atractiva e mais efectiva²³.

1.3.2. Terminologia e conceptualização

Sousa *et al.* (2012a: 27), que enquadram o desenvolvimento da conceptualização num ambiente de partilha e negociação e numa fase anterior à da especificação formal da ontologia, dividem, o processo de desenvolvimento de uma conceptualização em quatro fases principais: “*elicitação de conceitos, organização de conceitos, partilha de conceitos e negociação dos conceitos*”.

²³ Uma representação semiformal de um sistema conceptual oferece perspectivas de grande valor e interesse, partindo do princípio que o modelo conceptual definido é correctamente avaliado e validado pelos especialistas. O estudo dos problemas interlinguísticos a partir da sua representação conceptual é, a nosso ver, um dos exemplos dessa vitalidade.

Para o desenvolvimento e construção da conceptualização os autores vêm como essencial a conjugação de contributos de dois tipos de processos – *o terminológico e o da representação do conhecimento* – perspectiva que também defendemos e que está na base do método proposto nesta dissertação, que descreveremos no terceiro capítulo.

O trabalho terminológico engloba métodos para a identificação, selecção e classificação de recursos linguísticos, bem como técnicas de extracção terminológica. Desempenha também um papel de apoio às actividades de negociação que decorrem durante a análise e discussão sobre o conceito. O processo de representação do conhecimento, por sua vez, engloba as actividades de eliciação, organização, partilha e discussão das representações conceptuais, sendo que a actividade de eliciação dos conceitos pode, na opinião dos autores, ser também apoiada em processos terminológicos, de modo a ultrapassar as dificuldades inerentes a este processo (ao, por exemplo, permitir o acesso aos contextos de uso dos termos).

A terminologia é uma disciplina relacionada com o estudo do conhecimento, cujo objecto de estudo são os conceitos e as relações que estabelecem entre si, e a relação entre os conceitos e as suas designações, os termos. Tendo por base esta perspectiva, e dado que uma conceptualização construída colaborativamente é desenvolvida a partir da análise dos conceitos e dos termos que os designam, tanto os especialistas de domínio envolvidos no processo colaborativo como os terminólogos focalizam a sua atenção no mesmo objecto – o conceito. No entanto, como salientam Barros *et al.* (2012: 07),

while the former use terms and concepts for communicative and knowledge sharing purposes the latter study them in order to facilitate communication between experts in specialized domains or to enhance interoperability between information systems.

Esta dicotomia de análises e propósitos implica, assim, que os métodos terminológicos se adaptem ao ambiente comunicativo específico e às aplicações que o enquadram de modo a apoiar e a potenciar a fase de conceptualização de ontologias.

Como chama a atenção Santos (2010: 62) o suporte texto *“encerra conhecimento sob forma de manifestações linguísticas, mas o discurso sobre o mundo não se pode confundir com o mundo”*. Não se trata, para a autora, de minimizar a importância do discurso em língua natural. Trata-se antes de *“optimizar as potencialidades que esse discurso pode trazer para a codificação de representações do conhecimento”*.

Santos (2010: 134) explicita, por isso, a sua concordância com a opinião de Soares e Pereira (2008)²⁴ ao afirmar que as unidades linguísticas “*são pontos de acesso a repositórios mais vastos de conhecimento*”, e por isso um bom ponto de partida para determinado tipo de construção de conceptualizações, uma vez que “*os termos e as relações semânticas entre termos podem apontar para aproximações, maiores ou menores, a uma organização conceptual do domínio*”.

Do seu ponto de vista, embora uma abordagem semasiológica não seja adequada para, sozinha, formalizar conhecimento num sistema computacional, pode contribuir com informação sobre os conceitos e as suas designações, continuando o termo a desempenhar um papel fundamental na organização do conhecimento.

Esta interligação entre a terminologia e a representação do conhecimento no contexto da construção da conceptualização de um domínio é essencial, uma vez que, neste contexto, as actividades que envolvem a conceptualização requerem uma constante interpenetração entre a terminologia e a representação do conhecimento que seja capaz de:

rendering intuitive and operational the notions of term and concept without blurring the theoretical distinction between the different levels of analysis triggered by them. (Barros et al., 2012: 07)

No entanto, como advertem os autores, o trabalho prático nesta área, de que é exemplo a construção de ontologias, tende a colocar a representação do conhecimento e a terminologia como elementos alternados (por vezes opostos) ao invés de os perceber como elementos interdependentes de uma relação entre objectos, conceitos e termos, tal como representado pelo triângulo semiótico à luz da pesquisa em terminologia.

Roche & Costa (2011), no prefácio à conferência TOTH 2011²⁵, referem a importância desta interdependência e destacam o valor do uso dos princípios teóricos e metodológicos da terminologia para a construção de ontologias ao afirmarem que, no contexto das muitas aplicações informáticas baseadas em ontologias, quer sejam aplicações destinadas ao tratamento da informação ou à gestão do conhecimento, se coloca o problema:

²⁴ Para Pereira & Soares (2008: 03), que partem de um ponto de vista baseado na linguística cognitiva, a construção de significado “*is equated with conceptualization, a dynamic process where linguistic units serve as prompts for an array of conceptual operations and the recruitment of background knowledge*”.

²⁵ <http://www.porphyre.org/toth/toth-2011-en>

“de la prise en compte des terminologies lors de l’acquisition des connaissances du domaine. Si les corpus de spécialités constituent une source privilégiée d’information, il reste à définir des méthodes d’analyse qui tiennent compte des différences qui existent entre les textes et la conceptualisation formelle. La terminologie permet d’y répondre en distinguant les dimensions linguistique et conceptuelle. Elle établit de plus des principes théoriques et méthodologiques utiles pour la construction d’ontologies”.

Para Barros *et al.* (2012: 07), dado o panorama actual, o desafio coloca-se precisamente em manter a premissa dos *“termos como expressões linguísticas de unidades mentais e abstractas, os conceitos, no decurso do processo de conceptualização”*.

O uso de metodologias de base terminológica na fase de conceptualização, seja partindo de uma abordagem onomasiológica, pelo recurso à construção de representações conceptuais do domínio de especialidade com o auxílio do especialista, seja partindo de uma abordagem semasiológica, pelo recurso ao *corpus*, ou utilizando as duas abordagens de modo complementar²⁶, constituem, assim, a nosso ver, um contributo relevante no acesso à informação do domínio, promovendo a discussão e análise conceptual e contribuindo para obter melhores resultados na fase subsequente, a de formalização.

Tal como realçam Sousa *et al.* (2012), partindo de um ponto de vista prático:

From the creation and organization of a domain corpus, through term extraction methods for concept elicitation, until specific approaches to disclose semantic similarities during the discussion of achieved conceptual structures, terminology is a promising add-on of the process of conceptualization.

Entendemos, por isso, que uma abordagem interdisciplinar, que inclua a presença da terminologia e do terminólogo nesta fase, que não se limite apenas no trabalho de organização do conhecimento centrado na visão do especialista, e que conjugue, sempre que necessário, as perspectivas conceptual e linguística, pode conduzir a uma melhor elicitação e uma organização mais efectiva do conhecimento, produzindo um artefacto semântico mais rico, ponto de vista que procuraremos demonstrar na análise do nosso caso de estudo e na experimentação do método que propomos, descritas no capítulo quatro.

²⁶ Analisaremos com maior profundidade a relação entre estas duas perspectivas no ponto 1.4.1. e 1.4.2.

Por outro lado, o recurso às metodologias terminológicas no âmbito da conceptualização contribui, na nossa perspectiva, para apoiar o processo comunicativo que se desenrola entre os especialistas no trabalho colaborativo e, conseqüentemente, para promover uma gestão mais apoiada da dinâmica gerada em torno da elicitación, organização, partilha e negociação dos conceitos.

O conjunto das fases, actividades e recursos descritos implicam que dediquemos uma atenção cuidada à conceptualização de um domínio de especialidade, percebido enquanto processo de discussão iterativa e cíclica sobre o objecto, situado num contexto determinado e desenvolvido sob a influência de um conjunto de interlocutores e propósitos previamente definidos, processo que é partilhado, dinâmico e complexo e cujo resultado – uma representação conceptual – está sujeito à mudança e à evolução.

1.4. Ontologias

A conceptualização do mundo e a sua representação conduzem-nos ao conceito de ontologia (Roche, 2005: 52). Este conceito é um conceito multifacetado, utilizado por diferentes comunidades científicas com significados distintos. As visões mais distintas deste conceito emergem na comparação entre o que é uma ontologia do ponto de vista filosófico e do ponto de vista computacional.

Do ponto de vista filosófico a utilização primordial do termo ontologia deriva da Metafísica e de um dos tratados de Lógica de Aristóteles intitulado *Categorias*, obedecendo a um plano essencialmente didáctico de iniciação filosófica (Santos, 2010: 22). A Ontologia, enquanto disciplina, dedica-se ao estudo do *ser pelo ser* (Aristóteles, Meta. 1031a). Jacob Lorhard, reconhecido por ter cunhado o termo no seu livro *Ogdoas Scholastica*, define Ontologia como a ciência do que é inteligível pelo homem através da razão, sem influência da matéria.

*The science of the intelligible as intelligible insofar as it is intelligible by man by means of the natural light of reason without any concept of matter. (Lorhard, 1606: Livro 8, p. 1)*²⁷

²⁷ Tradução disponível em: <http://www.ilic.uva.nl/Publications/ResearchReports/X-2008-04.text.pdf>

Como descrevem Guarino *et al.* (2009: 01), a Ontologia²⁸, enquanto disciplina, dedica-se ao estudo:

*of attributes that belong to things because of their very nature. Unlike the experimental sciences, which aim at discovering and modeling reality under a certain perspective, Ontology focuses on the nature and structure of things per se, independently of any further considerations, and even independently of their actual existence.*²⁹

Do ponto de vista computacional, uma ontologia é vista como um tipo específico de objecto de informação ou artefacto computacional, representando este tipo de ontologia um meio para modelar formalmente a estrutura de um sistema, i.e., as entidades e relações relevantes que emergem da sua observação e que são uteis para um determinado propósito (Guarino *et al.*, 2009: 02)³⁰.

Segundo o filósofo Barry Smith (1998), uma ontologia é a ciência do que existe: *Ontology is the science of what is, of the kinds and structures of objects, properties, events, processes and relations in every area of reality.* Para o autor, no âmbito dos sistemas de informação, uma ontologia é uma representação de um domínio pré-existente da realidade que é inteligível ao especialista do domínio e que:

- *reflects the properties of the objects within its domain in such a way that there obtains a systematic correlation between reality and the representation itself,*
- *is formalized in a way that allows it to support automatic information processing*³¹.

Tendo por base esta perspectiva, uma ontologia, enquanto artefacto computacional, não tem como objectivo a compreensão do mundo, mas antes a representação dos objectos de modo a tornar possível a sua manipulação informática, pelo recurso a linguagens de

²⁸ Guarino *et al.* (2009) distinguem entre Ontologia e ontologia, atribuindo à primeira um carácter não-contável e à segunda um carácter contável.

²⁹ Não cabe no âmbito do nosso trabalho explorar detalhadamente o conceito de ontologia em termos filosóficos, pelo que esta análise será necessariamente breve e feita com o intuito de estabelecer um fio condutor que nos permita esclarecer a forma como o termo é usado mais recentemente noutras áreas disciplinares.

³⁰ Para os autores, o engenheiro de ontologias analisa as entidades relevantes (uma entidade denota “*the most general being, and, thus, subsumes subjects, objects, processes, ideas, etc.*”) e organiza-os em conceitos e relações, que são representados, respectivamente, por “*unary and binary predicates. The backbone of an ontology consists of a generalization/specialization hierarchy of concepts, i.e., a taxonomy*” (Guarino *et al.*, 2009: 02).

³¹ Buffalo Ontology Site: <http://ontology.buffalo.edu/>

representação formais (Roche, 2005: 53)³². Uma ontologia assume, assim, o papel de uma representação sistemática da realidade tal como é percebida, sendo representada através de uma linguagem formal específica usada para descrever uma determinada realidade, acrescida de um conjunto de decisões explícitas, de modo a fixar de forma rigorosa o significado pretendido.

Para Sowa (2006), a ontologia dedica-se ao estudo das categorias das coisas que existem ou possam existir num dado domínio, representando uma ontologia um catálogo do tipo de coisas que se assume que existem num *“domain of interest D from the perspective of a person who uses a language L for the purpose of talking about D”*.

De acordo com Gruber (1993; 1993a), no caso da IA a abordagem ao que existe é pragmática, i.e., nos sistemas de IA o que existe é aquilo que pode ser representado. Uma ontologia corresponde, na sua definição bem conhecida, a uma *especificação explícita de uma conceptualização* (Gruber, 1993). Gruber (1993a) sustenta do seguinte modo a sua definição:

In the context of knowledge sharing, I use the term ontology to mean a specification of a conceptualization. That is, an ontology is a description (like a formal specification of a program) of the concepts and relationships that can exist for an agent or community of agents. This definition is consistent with the usage of ontology as a set-of-concept definitions, but more general. And it is certainly a different sense of the word than its use in philosophy.

Formalmente uma ontologia é uma declaração de uma teoria lógica definindo um conjunto de primitivos de representação com os quais modelamos o domínio do conhecimento ou do discurso, sendo esses primitivos constituídos por classes, atributos e relações, como sintetiza Santos (2010: 29), com base em Gruber (1993; 2008). As definições incluem informação sobre o seu significado, enquanto os axiomas formais restringem a interpretação e o uso bem formado dos termos. Por outro lado, as ontologias, enquanto especificações, resultam de uma soma de partes formais e informais. Segundo a autora, as partes informais ajudam a explicar alguma coisa aos humanos, as partes formais permitem a análise automática.

³² Roche (2003: 02) vê uso do termo ontologia como uma escolha desafortunada, dado que possui um significado bem definido na literatura filosófica pouco relacionado com a descrição de conteúdos em repositórios de informação.

Esta definição sofreu uma evolução importante quando Borst (1997) a redefiniu como uma *especificação formal de uma conceptualização* partilhada, tendo introduzido o conceito de partilha, que implica negociação e consenso entre o conjunto dos interlocutores envolvidos na construção de uma ontologia. Esta definição foi depois conjugada com a anterior proposta por Struder *et al.* (1998: 161), que definiram uma ontologia como uma *“especificação formal e explícita de uma conceptualização partilhada”*³³. Staab & Struder (2009: viii) decompõem e analisam os componentes desta definição do seguinte modo:

‘Explicit’ refers to the fact that all elements of an ontology are explicitly defined, whereas ‘formal’ means that the ontology specification is given in a language that comes with a formal syntax and semantics, thus resulting in machine executable and machine interpretable ontology descriptions. Finally, ‘shared’ captures the aspect that an ontology is representing consensual knowledge that has been agreed on by a group of people, typically as a result of a social process.

As ontologias resultam, de acordo com a perspectiva descrita, de um processo social e representam formas apriorísticas de atingir um acordo sobre uma conceptualização relativa a uma dada parte do mundo, pelo que obter um consenso à volta da conceptualização representa um passo fundamental para o sucesso da ontologia (Pereira & Soares, 2008). Tal como descrevem os autores, uma metodologia de desenvolvimento de ontologias deve ser composta por um conjunto de princípios, processos, práticas, métodos e actividades bem estabelecidos, utilizadas na sua construção, avaliação e implementação.

Tendo por base a análise de Corcho (2003: 53) quando conclui que o objectivos das ontologias é o de capturar *“consensual knowledge in a generic and formal way, and that they may be reused and shared across applications (software) and by groups of people. Ontologies are usually built cooperatively by a group of people in different locations.”*, Pereira & Soares (2008) destacam a importância de incluir os princípios de colaboração no processo de desenvolvimento de ontologias, sobretudo na fase de especificação.

Para os autores, um processo de especificação colaborativa de uma ontologia corresponde a um conjunto de práticas e actividades utilizadas para obter uma conceptualização partilhada do domínio com a participação de todas as partes interessadas. Compreende a identificação dos conceitos a incluir na ontologia, as suas características,

³³ Nosso sublinhado.

definição e relações, a organização e a estruturação do conhecimento com recurso a representações externas independentes da linguagem e do ambiente de implementação e contém os seguintes princípios:

1. participação activa de todas as partes; o processo requer a colaboração constante entre a equipa de desenvolvimento e os restantes parceiros/especialistas;
2. existência de métodos que apoiem o processo de negociação de modo efectivo e eficiente, bem como de métodos que apoiem a obtenção de acordos e de consenso sobre o conteúdo da ontologia entre grupos de actores humanos;
3. existência de mecanismos que possibilitem o acesso e o trabalho com base nas diferentes perspectivas apresentadas pelos diferentes utilizadores;
4. existência de ferramentas que apoiem a colaboração (comunicação, cooperação e coordenação), baseadas, por exemplo, no recurso à visualização gráfica dos conteúdos propostos para a conceptualização partilhada durante o processo de análise e negociação do conhecimento do domínio;
5. existência de técnicas que apoiem a elicitação de conceitos/termos;
6. existência de mecanismos que apoiem a análise semântica e sintáctica dos conceitos da ontologia, de modo a garantir uma interpretação correcta dos diferentes contributos recebidos (Pereira & Soares, 2008).

Estes princípios são, na nossa perspectiva, fundamentais para um desenvolvimento efectivamente partilhado da ontologia e para uma representação mais granular do conhecimento do domínio.

Uma ontologia representa, então, um modelo de conhecimento de um domínio particular desenvolvido num dado momento no tempo, modelo partilhado entre dois ou mais actores do domínio (Buitelaar *et al.*, 2006: 503), com base em três dimensões essenciais: *“o conhecimento (uma conceptualização para entendermos o mundo), a língua (para falarmos sobre o mundo) e a lógica (para representar a manipulação do nosso entendimento sobre o mundo)”* (Santos, 2010: 38).

Como afirma Lenci (2008) o conhecimento é criado ou adquirido com um propósito, de modo a ser usado como uma ferramenta para atingir um determinado objectivo ou realizar uma tarefa específica. Mas, o seu uso resulta, também, segundo o autor, na

modificação do conhecimento prévio sobre as entidades e processos, levando-nos a rever e a actualizar a ontologia.

Do mesmo modo, o emprego de um elemento de conhecimento para a realização de uma tarefa pode produzir conhecimento novo que, por seu turno, pode ser acrescentado à ontologia, o que conduziria a uma revisão mais ou menos profunda da sua estrutura. Para Lenci (2008) este é, normalmente, o propósito principal em que assenta a realização da actividade de construção de uma ontologia.

Buitelaar *et al.* (2006: 503) defendem uma perspectiva próxima da de Lenci. Segundo os autores, a definição de uma ontologia depende primariamente do conteúdo que já foi alvo de interpretação. Tendo em atenção este facto, Buitelaar *et al.* afirmam que:

content production and interpretation will drive the adaptation of the ontology infrastructure, and ontology adaptation will drive content interpretation and production. In order to arrive at such a continuous 'hermeneutic cycle' of content and knowledge production and interpretation, a rich representation of domain knowledge and content features is needed.

O desenvolvimento deste ciclo contínuo permitirá, a nosso ver, que uma ontologia possa atingir mais efectivamente a sua finalidade: a de ajudar a compreender o mundo a partir da representação de conhecimento sobre esse mundo ou sobre uma parte dele.

No entanto, dado que, para os sistemas de IA, o que existe é o que pode ser representado numa linguagem formal - objectos, conceitos e relações que se representam de forma explícita através da ontologia - somos levados a concluir, tal como Santos (2010: 32), que embora uma ontologia seja um *“artefacto de engenharia que pode ser partilhado entre pessoas, organizações e sistemas informáticos, a representação sistemática da realidade não é garantia de amplitude de conhecimento, se o que existe é somente o que pode ser representado formalmente”*.

De facto, e como reconhece a autora, a popularidade das ontologias enquanto instrumento de resolução para os problemas de organização e representação do conhecimento exige propriedades que são difíceis de reunir num mesmo contexto. Para Santos (2010: 38,39), podemos normalizar linguagens de representação e poderíamos, hipoteticamente, normalizar a língua; o que não podemos é normalizar o pensamento, pelo

que afirma que uma “*ontologia computacional é a especificação de uma conceptualização dos objectos de um domínio segundo um determinado ponto de vista imposto pela aplicação*”, e, no nosso ponto de vista, imposto também pelos propósitos e pelo grupo de especialistas que orientam a sua criação.

1.4.1. Terminologia e ontologias

Terminologie et ontologie. Deux termes que pourraient s’opposer à plus d’un titre – le premier portant sur une pratique, le deuxième relevant de la métaphysique – et qui sont néanmoins de plus en plus associés. (Roche, 2005: 48)

A construção de um sistema que represente o conhecimento de um determinado domínio obriga, tal como defende Costa (2006: 9), em concordância com Rastier (1995: 43-44), a agir simultaneamente sobre:

o estado das coisas, as redes conceptuais e as estruturas semânticas, sendo que, do ponto de vista da análise, estes representam três níveis diferentes de representação do conhecimento que, pelo recurso a modelos formais, nos permitem activar sistemas computacionais ao interligar o pensamento, a língua e as leis da natureza.

Remígio (2010: 25) chama a atenção para o facto, inevitável, deste processo de representação do conhecimento não ser uma tarefa simples, nem encontrar uma forma única possível. Por um lado, porque no acto de conhecer o real e de organização e sistematização desse conhecimento,

o sujeito tem que considerar a constante presença da desordem e de si próprio no mundo que o rodeia e, por outro, porque a estruturação do conhecimento se baseia num consenso por parte da comunidade de especialistas que o produz, estruturação que, apesar de colectiva, tem igualmente uma vertente subjectiva, dado que cada especialista possui uma representação mental própria.

A organização do conhecimento implica, assim, a existência de estruturas conceptuais que representem e tornem claras as relações que se estabelecem entre os conceitos de um domínio. Tal como explicita Costa (2009: 03) a criação e a produção

científica implicam “a sistematização, a estruturação e, eventualmente, a hierarquização do conhecimento, sendo estas actividades o resultado da conceptualização dos referentes e dos objectos”.

A análise e validação da pertença dos conceitos a domínios de especialidade e a sua estruturação em sistemas de representação de conhecimento especializado constituem princípios e tarefas fundamentais da terminologia. Para Cabré (1998: 167) o conceito surge como o elemento que estabelece a ligação entre a terminologia e a teoria do conhecimento. Para a autora, o conjunto dos termos de um domínio de especialidade constitui, ao mesmo tempo, uma estrutura de conceitos que reflecte a organização do conhecimento sobre o domínio em questão. Os termos tornam-se, deste modo, elementos chave na representação do conhecimento especializado, dado que um sistema de gestão da informação, ao processar a informação, internaliza o conhecimento estruturado com recurso aos termos. Cada termo é um item de pensamento especializado e reflecte a organização conceptual do domínio de especialidade.

Para proceder à organização do conhecimento, torna-se, assim, essencial, na nossa opinião, recorrer à terminologia. O recurso a esta disciplina científica é, na nossa perspectiva, e recordando as palavras de Roche (2008: 55), indispensável, tendo em conta que o seu objecto central é o de compreender o mundo, descrever os objectos que o povoam e encontrar “*les mots justes pour en parler*”.

A terminologia desempenha, segundo o autor, alternadamente, dois papéis – o de uma prática e o de uma ciência. Recorre à linguística para a “*prise en compte des langues d’usage et s’en distingue par la modélisation scientifique de la signification des mots*”. Recorre à epistemologia para a pesquisa de uma compreensão do mundo e distingue-se pela importância que atribui aos termos, combinações indissociáveis de uma designação e de um conceito (Roche, 2005: 54).

Entendemos, tal como Roche (2008), que nem uma ontologia é uma terminologia, nem uma terminologia é uma ontologia³⁴, uma vez que o seu objectivo principal não é o de especificar formalmente uma conceptualização. No entanto, podem estar intrinsecamente interligadas, como o destacam os trabalhos e abordagens propostas por Bourigault e

³⁴ Ainda que certas definições levem a pensar o contrário: “*a set of definitions of a formal vocabulary*” (Gruber, 1992).

Aussenac-Gilles (2003), Temmerman (2005), Budin (2005), Barros *et al.* (2012) ou pelo próprio Roche (2006b, 2007b, 2009), dada a proximidade entre metodologias e propósitos.

Como indica Ceusters (2008), a terminologia, quando perspectivada do ponto de vista da construção de ontologias, não representa aquilo a que chama o meio apropriado (*the right device*) para:

- *explain why reality is what it is, how it is organised, etc., (although it is needed to allow communication),*
- *reason about reality,*
- *make machines understand what is real,*
- *integrate across different views, languages, conceptualisations, ...*

Para o autor a terminologia procura, sobretudo, melhorar as condições de comunicação entre humanos. Entendemos, no entanto, que esta visão é limitativa. No nosso ponto de vista, como uma disciplina interdisciplinar com as suas origens na comunicação especializada, a terminologia encerra aspectos de grande sinergia e interligação com a engenharia do conhecimento e com o desenvolvimento de sistemas de gestão do conhecimento baseados em ontologias.

De facto, se entendidas enquanto produtos informáticos com uma finalidade específica de aplicação, exige-se que tanto ontologias, como terminologias, possuam as mesmas propriedades de consenso, coerência, precisão e de capacidade de evolução de modo a poderem ser aceites e (re)utilizadas (Roche, 2005: 51)³⁵.

Como afirma Santos (2010: 94), é fundamental percebermos que a terminologia no seu momento actual encontra-se profundamente ligada à inovação em todas as suas vertentes: científica e académica, social, cultural e económica. Para a autora,

A terminologia e a ontologia servem propósitos utilitários, mas o conceito de conhecimento para um terminólogo quase sempre tem contornos diferentes do conceito de conhecimento para um engenheiro do conhecimento, pelo que se torna essencial reduzir esse afastamento para que ambas as práticas possam colaborar eficazmente.

³⁵ Roche (2005: 51) considera que, para além de um processo, uma terminologia é também um produto, resultante de uma prática industrial devendo, para tal, satisfazer um conjunto de propriedades necessárias à sua reutilização, como ser consensual, coerente, precisa nas suas definições e evolutiva nos seus sistemas nocional e de designações.

Para a autora, este encontro de sinergias deve procurar extrair o melhor de dois mundos: por um lado, a normalização através da estabilização dos significados em representações formalizadas, por outro, a captura da exuberância criativa e ininterrupta da língua e do pensamento humano. Esta interligação implica, de acordo com Gresser (2008), uma cooperação estreita entre especialistas, terminólogos e engenheiros de ontologias para detectarem atempadamente a emergência de novos conceitos.

Em comum terminologias e ontologias partilham raízes e propósitos. Como identifica (Budin, 2005), a sua raiz comum reside na lógica clássica, com as categorias Aristotélicas e as hierarquias conceptuais. Partilham objectivos e funções, quer na organização de dados, informação, conhecimento, conteúdos; quer no apoio à tradução automática e à tradução assistida.

Partilham também o facto de servirem um propósito idêntico: o de fornecer uma organização partilhada do conhecimento sobre uma parte específica do mundo a diferentes utilizadores, de modo a facilitar o acesso àquele e a promover uma comunicação eficiente. Por outro lado, ambas as disciplinas organizam e representam o conhecimento com base em sistemas conceptuais, representação que procura ser independente de uma qualquer língua natural.

Terminologias e ontologias procuram, assim, expressar e explicitar o conhecimento e possuem como objectivo alcançar uma representação comum de um domínio de especialidade, fornecendo meios e metodologias para auxiliar as pessoas a partilhar conhecimento dentro de uma organização ou em comunidades profissionais distribuídas que trabalham para atingir objectivos comuns.

Possuem, necessariamente, e como explicitámos antes, diferenças entre si, nomeadamente, ao nível do grau de formalização e explicitação. Para Budin (2005), partindo de um ponto de vista operacional, a diferença, que descreve como gradual, entre os graus de formalização e de explicitação tem implicações importantes, uma vez que implica o dispêndio de um grande esforço intelectual na transformação do conhecimento implícito em explícito e na sua formalização.

A este esforço acrescentam os limites temporais normalmente impostos à construção de uma ontologia, a complexidade do processo de construção de representações de

conhecimento, bem como a necessidade de encontrar e validar fontes que possibilitem o acesso à informação e ao conhecimento sobre o domínio, factores que, no seu conjunto, constituem restrições relevantes ao processo aquisição e representação do conhecimento de domínio com base numa ontologia. Tendo em conta estes factores, a terminologia surge, também aqui, como um elemento que pode desempenhar um papel chave na diminuição destas restrições, sobretudo no que respeita à análise do conhecimento do domínio, a partir, por exemplo, do recurso aos *corpora* e à análise, estruturação e representação dos dados advindos do seu processamento.

Verificamos, na anterior afirmação de Budin, uma outra diferença, ao nível dos utilizadores, dado que uma terminologia, enquanto conjunto organizado de termos de um domínio, se destina sobretudo a utilizadores humanos e as ontologias são desenvolvidas para a partilha de conhecimento entre agentes (tanto humanos, como artificiais).

Como descrevem Gamper *et al.* (1999), de modo a comunicar sobre uma parte do mundo real, o conhecimento sobre essa parte tem que ser representado por um qualquer tipo de modelo de conhecimento com o qual concordem todos os agentes que participam do processo. O principal objectivo destes modelos reside na facilitação de uma comunicação efectiva entre os vários agentes através da disponibilização de um método para uma representação concisa e sem ambiguidades do conhecimento a ser transferido³⁶, de modo a que dois agentes *“either humans or artifacts such as computer systems, are able to communicate and possibly argue about the part of the world described by the model”*.

Também ao nível da definição existem diferenças, uma vez que a terminologia recorre ao texto em língua natural para definir o significado dos termos, enquanto uma ontologia especifica conhecimento de modo explícito com recurso a uma linguagem formal com uma semântica bem definida, de modo a permitir uma interpretação do termo livre de ambiguidades (Gamper *et al.*, 1999).

³⁶ Como reconhecem Buitelar *et al.* (2011: 33), a língua natural, é, muitas vezes, o meio escolhido para a representação e transferência de conhecimento entre humanos. No entanto, a ambiguidade é um elemento inerente à língua natural. *“Words have multiple meanings and grammar can be ambiguous in structure and therefore in interpretation”*. E, ainda que o autor reconheça que estas ambiguidades se resolvem com alguma facilidade entre interlocutores humanos, tal já não acontece quando o interlocutor é a máquina e o nosso propósito o da construção de uma ontologia.

No entanto, como afirma Budin (2005: 106), as terminologias estão limitadas aos conceitos e às suas designações no âmbito de um domínio específico, limitação que não afecta as ontologias, recaindo o seu enfoque, usualmente, nas suas funções comunicativas e linguísticas.

Por outro lado, constituindo as ontologias um meio para operacionalizar modelos conceptuais que representam o conhecimento do domínio, torna-se, no seu desenvolvimento, necessário utilizar e integrar métodos de modelação conceptual focalizados naquele conhecimento, destacando-se neste processo, a nosso ver, e como apontámos antes, as abordagens propostas pela terminologia à conceptualização do domínio, uma vez que se dedicam, com grande especificidade, à análise da natureza do conhecimento do domínio, i.e., da sua estrutura conceptual.

A construção de estruturas conceptuais representativas do conhecimento do domínio são, há muito, prática corrente no trabalho e na pesquisa terminológica, pelo que os modelos terminológicos referentes ao conhecimento do domínio são uteis para explorar a possibilidade de transformar estas estruturas em ontologias. Na perspectiva de Budin (2005: 105):

The bi-functionality of terminology modeling is crucial for ontology engineering because terminological knowledge models may represent both, the knowledge expressed in a particular text or in a whole collection of interrelated texts, as well as the knowledge of a whole domain, by exploiting a multitude of texts by extracting specific concepts from them and building a domain-specific concept system.

Concordamos também com o autor quando afirma que o recurso a abordagens terminológicas pode contribuir para aumentar a granularidade no processo de explicitação das estruturas conceptuais, permitindo, em particular, que se clarifiquem as diferenças culturais dessas estruturas, tal como são expressas pelas diferentes comunidades linguísticas em respeito pelas suas convenções e valores que contribuem para dar forma aos seus discursos de domínio (Budin, 2005: 120).

Defendemos esta perspectiva uma vez que entendemos, tal como Costa (2005), que a construção de uma ontologia é sobretudo uma tarefa que requer uma

“abordagem interdisciplinar³⁷ que permita, a partir de um ponto de vista determinado por imperativos sociais, culturais, políticos e/ou económicos previamente identificados, uma descrição mais completa, equitativa e fiel da realidade a modelizar, constituindo uma ontologia um recurso privilegiado para a representação do conhecimento”.

Para que essa descrição da realidade se torne viável, devemos, a nosso ver, atender com cuidado aos requisitos necessários ao desenvolvimento de uma ontologia, procurando interligar com equilíbrio os (1) processos e metodologias definidos para apoiar a sua construção, (2) o grupo de especialistas de domínio reunido para a sua construção e validação e (3) as aplicações a utilizar para apoiar a conceptualização (partilhada) do conhecimento domínio e a sua formalização.

1.5. Terminologia e organização do conhecimento

The major challenge of terminology is conceptual, not linguistic: terminologists are trained in linguistics and thus are properly prepared for the linguistic dimension of their task; in contrast, they are not normally domain experts, yet they require a substantial amount of expert knowledge in order to do their work. In other words, the major difficulty is pinning down the meanings of terms. (Meyer, 1992:23)

Tal como já referimos, a terminologia tem por objecto de estudo os conceitos e as suas designações, os termos, bem como as relações que aqueles estabelecem entre si, no seio de um domínio de especialidade, estando inextricavelmente ligada ao conceito de conhecimento especializado, à sua organização e à sua representação. Combina, para o desenvolvimento do trabalho terminológico, elementos de diferentes abordagens teóricas que se dedicam à descrição, organização e transferência do conhecimento.

Enquanto disciplina, a terminologia dedica-se, por um lado, ao estudo dos princípios e métodos que governam o estudo dos conceitos e das suas designações em domínios de especialidade e, por outro, às actividades de recolha, gestão e representação de informação relevante para o domínio. A interdisciplinaridade da terminologia resulta, por seu lado, do

³⁷ Nosso sublinhado.

carácter multifacetado dos termos, que podem ser analisados enquanto elementos linguísticos, elementos conceptuais e elementos de comunicação. Sager (1990: 13) distingue três dimensões do termo – cognitiva, linguística e comunicativa - e advoga, em consequência, uma teoria da terminologia que englobe essas três dimensões:

- *a cognitive one which relates the linguistic forms to their conceptual content;*
- *a linguistic one which examines the existing and the potential forms of the representation of terminologies;*
- *a communicative one which looks at the uses of terminology and has to justify the human activity of terminology compilation and processing.*

Também Cabré defende, em diferentes momentos (1999; 1999a; 1999b; 2003), esta natureza multidimensional dos termos. Para a autora, ao aceitarmos a natureza multidimensional das unidades terminológicas estamos perante três dimensões (*cognitiva, linguística e comunicativa*) que devemos ter sempre presentes como ponto de partida, dado que cada uma delas, apesar de constituir um elemento inseparável da unidade terminológica, permite por si só um acesso directo ao objecto (2003: 187).

Para Rey (1995), a dimensão cognitiva da terminologia resulta da natureza cognitiva do seu objecto de estudo, não sendo esta uma particularidade respeitante apenas à sua teoria. O que se torna interessante e relevante na terminologia é o modo como estabelece com sucesso a ligação entre a dimensão cognitiva e as necessidades sociais por via das formas linguísticas, contribuindo assim para a optimização da comunicação entre especialistas e outros mediadores que transferem conhecimento, como terminólogos, tradutores, gestores de conhecimento, entre outros, e o público em geral.

De acordo com Sager (1990: 13), abordar o estudo da terminologia a partir da sua dimensão cognitiva requer uma compreensão da estrutura do conhecimento de modo a obter uma imagem tão completa e coerente quanto possível da natureza, comportamento e interacção entre os conceitos e termos a eles associados. A dimensão cognitiva tem por fim a análise da relação entre conceitos e do modo como se organizam em sistemas conceptuais que representam um dado domínio do conhecimento.

Esta representação, desenvolvida com o objectivo de apoiar a transmissão de conhecimento com precisão e sistematicidade em diferentes situações, é dinâmica e está, por conseguinte, sujeita a reformulações, uma vez que reflecte, no momento da sua

construção, o estado actual do conhecimento, tal como é reconhecido e partilhado pelos membros de uma comunidade.

O carácter evolutivo das representações conceptuais, através das quais se desenvolve o processo de conceptualização, resulta do facto de uma representação conceptual de um domínio não poder ser entendida como um sistema totalmente fechado, uma vez que o conhecimento especializado é uma subdivisão do nosso sistema de conhecimento geral, sendo composto por esquemas cognitivos flexíveis que permitem, até certo ponto, a manipulação das ideias e a discussão e construção de hipóteses sobre o conhecimento representado ou a representar.

A possibilidade de evolução leva Bessé (1997: 67) a afirmar que as representações conceptuais não são absolutas e reflectem apenas uma interpretação particular da realidade, dado que nem mesmo as classificações mais rigorosas conseguem conter todos os conceitos constitutivos de um domínio específico. Esta perspectiva está de acordo com os pontos de vista defendidos por Sager (1990) e Cabré (1999, 2004), para quem, no desenvolvimento de um sistema de representação do conhecimento, a representação conceptual é construída de acordo com as necessidades e propósitos dos seus construtores e/ou dos seus potenciais utilizadores, tendo em conta os valores sociais e culturais vigentes.

Tanto a dimensão do sistema conceptual, como a complexidade das relações a serem estabelecidas no seu interior resultam de uma avaliação prática do propósito que subjaz à realização da análise conceptual. Como defende Cabré (2004: 111) existe mais do que uma possibilidade de representação conceptual de cada domínio, se considerarmos a poliedricidade do conhecimento e o seu carácter dinâmico, tanto do ponto de vista diacrónico, como sincrónico.

O desenvolvimento do trabalho terminológico pode ter por base pontos de partida distintos, assente em teorias terminológicas distintas, que operacionalizam as metodologias de acordo com um conjunto de objectivos pré-estabelecidos. Como refere Wright (2007: 157), qualquer análise abrangente das práticas terminológicas correntes, deve dar conta da tensão fundamental entre:

the notion of linguistic indeterminacy and the impetus to model knowledge based on semantic ordering principles that are nonetheless inextricably represented by language, with the ever present potential for ambiguity.

Tal como explicitamos antes, as metodologias aplicadas no trabalho terminológico e na investigação em terminologia estão, normalmente, ligadas a uma de duas perspectivas: a perspectiva conceptual, que remete para uma abordagem onomasiológica, e a perspectiva linguística, que remete para uma abordagem semasiológica.

O trabalho terminológico, como descreve Costa (2010), quer seja levado a cabo por terminólogos ou por engenheiros do conhecimento focaliza-se prioritariamente em uma das facetas do termo:

les premiers valorisent tendanciellement l'expression linguistique du concept, la dénomination, les deuxièmes se concentrent préférentiellement sur l'entité dénommée, le concept.

A terminologia que parte de uma perspectiva linguística centra, assim, a sua atenção na verbalização do conhecimento, partindo do termo para o conceito, fundamentalmente através da análise do texto. A terminologia que parte de uma perspectiva extralinguística focaliza-se, por seu lado, na conceptualização do conhecimento, partindo do conceito para o termo, e tem em conta sobretudo a contribuição do especialista de domínio. A decisão de recorrer a uma ou outra das perspectivas acima identificadas tem consequências a diversos níveis da análise e organização do conhecimento, como procuraremos explicitar em seguida.

1.5.1. Perspectiva linguística

Partindo de uma perspectiva linguística, o objecto da terminologia centra-se essencialmente no termo. Como define a norma ISO 704 (2009: 34) – o termo é uma “designação que consiste em uma ou mais palavras e que representa um conceito geral numa língua de especialidade de um domínio específico”. Partindo de uma perspectiva linguística, de base semasiológica, um termo é uma entidade lexical utilizada num determinado domínio de especialidade e, por isso, um signo linguístico³⁸.

³⁸ Esta perspectiva, em que a unidade terminológica é vista como um signo advém, segundo Costa (2006: 79) da concepção de signo linguístico apresentada por Cassirer (1972: 27): « *le signe n'est pas l'enveloppe qu'un pur hasard attribuerait à la pensée, mais son organe nécessaire et essentiel. Il ne sert pas seulement aux fins de la communication d'un contenu de pensée qui serait donné une fois achevées: il est un instrument au moyen duquel ce contenu lui-même prend forme en s'extériorisant et au moyen duquel seulement il acquiert*

Como explicita Prandi (2011), no caso de um signo pertencente a uma língua natural, o que perfaz a relação biunívoca entre significante e significado é a partilha entre uma comunidade de locutores; no caso de um termo, o que perfaz esta mesma relação é a partilha entre um grupo de especialistas que, por um lado, não é mais do que um conjunto de elementos de uma comunidade linguística, mas que, por outro, idealmente, intercepta diferentes comunidades linguísticas.

Para Lerat (2009), o termo pode ser visto como o nome dado numa língua a uma entidade conceptualizada por uma comunidade de trabalho. Os termos constituem, assim, unidades lexicais de especialidade, pelo facto de representarem um conhecimento específico a uma área do saber, reconhecido e partilhado pelos membros de uma comunidade de especialistas. Esse conhecimento é o resultado do domínio de um saber, que se traduz no conhecimento que se possui das relações estabelecidas entre os termos, que num contexto textual e/ou discursivo denominam as relações que se estabelecem entre conceitos num dado sistema conceptual.

Ao centrar-se nos termos e nas relações entre os termos tal como se materializam no texto, a perspectiva linguística vê o texto como o veículo privilegiado de manifestações linguísticas que podem reflectir uma organização conceptual, tal como descreve Pearson (1998: 02), quando afirma que os terminólogos, independentemente do seu posicionamento “tradicional” ou “moderno”, se preocupam, essencialmente, com a construção de estruturas de conhecimento de domínios específicos, estruturas que constroem a partir de um processo de comparação e contraste entre conceitos relacionados. Para a autora,

“The ordering and classification of knowledge are crucial to terminological studies. In ‘modern’ terminology practice, the emphasis is much more on usage with the use of ‘real text as primary source of data’”.

Numa perspectiva linguística, a terminologia centra-se em áreas de especialidade nas quais estuda o conhecimento na sua relação com a expressão, interessando-se, como refere Santos (2010: 87), tradicionalmente e em particular, pela parte linguística através da qual o conceito é designado ou descrito e pelas relações que se estabelecem entre esses conceitos,

la plénitude de son sens. La détermination conceptuelle d’un contenu va de pair avec la fixation dans un signe caractéristique ».

nomeadamente através das relações semânticas de hiperonímia, hiponímia ou de meronímia.

Uma concepção linguística da terminologia implica, segundo Cabré (2003: 186), reconhecer que os objectos de conhecimento são as unidades terminológicas concebidas como combinações indivisíveis de forma e conteúdo. Para a autora, o conteúdo está associado à forma e *“therefore the units not only designate but they also mean with all the cognitive consequences of their meaning”*.

Estas unidades, que define como combinações de condições linguísticas, aparecem no discurso especializado, discurso que, na assunção da autora, apresenta uma estrutura organizada do conhecimento, que pode resultar numa representação através de um mapa conceptual composto por *“nodes of knowledge, which can be represented by different types of units of expression, and by relations between these nodes.”* (Cabré, 2003: 189)

Para Conceição (2005: 67), que parte de uma perspectiva linguística, a organização conceptual constrói-se através da língua, partindo-se do discurso para comunicar. Um termo serve, assim, e ao mesmo tempo de veículo de transmissão e compreensão do conhecimento, uma vez que, enquanto unidade linguística é, na sua essência, plural. Como descreve o autor, recorrendo a Langacker (1997: 24), um termo é uma *“creation and reflection of a culture as well as a primary instrument of its constitution and transmission”*. Os termos são, para o autor, unidades lexicais especializadas que funcionam num sistema linguístico próprio de cada utilizador e de cada comunidade, sendo que o estudo dos termos deve ocorrer dentro das diferentes situações comunicativas.

Na perspectiva de Conceição (2005: 69) a pesquisa terminológica deve, primeiro, concentrar-se numa análise exaustiva dos termos a partir do texto ou de corpora textuais e só em seguida partir para a análise do nível conceptual e para a consulta ao especialista de domínio. Na sua opinião, os termos representam manifestações de conceptualizações que:

“ne peuvent être décrites par le terminologue que par le travail sur les discours et leur fonctionnement. L’analyse terminologique ne s’occupe pas seulement, dans ce cas, de la recherche de la dénomination du concept, mais aussi de l’analyse des occurrences discursifs des dénominations que aide à délimiter les concepts dénommés et comprendre les savoirs exprimés. (...) Les termes son envisagés en tant qu’unités de connaissance, accessibles à partir de l’existence discursive.” (Conceição, 2005: 69-70)

Na abordagem linguística, o sentido de uma unidade lexical é, assim, definido ao observar o conjunto das suas interações com outras unidades linguísticas num determinado contexto sendo que a observação das oposições entre essas unidades, as similitudes e a análise das suas combinatórias permite circunscrever gradualmente o seu sentido (Santos, 2010: 78).

Nesta perspectiva, é na relação entre o saber linguístico e o saber específico de um domínio de conhecimento que reside o aspecto interdisciplinar da pesquisa terminológica, pesquisa que, no entanto, deve ser levada a cabo em articulação com os especialistas do domínio (Conceição, 2005: 28)³⁹.

No caso da construção de ontologias, as metodologias que assumem uma perspectiva linguística partem do pressuposto que é possível construir ontologias a partir de textos. Para levar a cabo a sua construção baseiam-se essencialmente em corpora textuais, submetendo-os a análises estatísticas, semânticas e lexicais, para tentar extrair candidatos a termo e relações⁴⁰ entre esses termos que denunciem uma eventual estrutura conceptual que, por sua vez, seja a tradução da organização de conhecimento.

Mondary *et al.* (2008: 87) descrevem esse processo como tendo, normalmente, quatro etapas: (1) *constituição de um corpus de documentos*, (2) *análise linguística do corpus*, (3) *conceptualização* e (4) *operacionalização da ontologia*, destacando a fase de conceptualização como uma fase crucial à construção de uma ontologia a partir de textos, uma vez que é nesta fase que se efectua a passagem do nível do discurso ao nível conceptual.

Segundo os autores, estas etapas, que vêm como relativamente independentes, dão origem a um duplo movimento que permite a passagem (1) do nível textual (conhecimento descrito nos textos) ao nível conceptual (o conhecimento é descrito por via dos conceitos denotados por entidades linguísticas e pelas relações entre esses conceitos) e (2) do informal ao formal.

³⁹ A abordagem seguida por Conceição está de acordo com os princípios propostos pela socioterminologia e pela teoria sociocognitiva, que partilham do ponto de vista que a pesquisa e o trabalho terminológico se devem basear em corpora textuais.

⁴⁰ A extracção de relações semânticas a partir de textos é descrita por Auger e Barrière (2008: 02) como um processo capaz de minimizar “*the labor-intensive phase of manual knowledge engineering and thus overcoming the long-standing knowledge acquisition bottleneck.*”

Passar dos textos a uma ontologia significa, nesta perspectiva, passar da realização linguística ao modelo conceptual. No entanto, e como reconhecem os autores, construir uma ontologia a partir do *corpus* oferece limitações e não pode ser um processo que decorra de forma inteiramente automática, existindo diferentes possibilidades quanto ao caminho a seguir na construção do sistema conceptual, dado que, na sua perspectiva,

Les textes ne désignent pas les concepts du domaine même si, à travers le vocabulaire employé, on voit poindre des notions importantes, même si on peut se faire une idée de l'envergure du domaine à modéliser et donc de la couverture de l'ontologie à produire, il reste un important travail de conceptualisation à faire (Mondary et al., 2008: 94).

O recurso à abordagem de base linguística é visto como útil e necessário no contexto do trabalho terminológico por diferentes autores (Temmerman, 2000; Rogers, 2004; Marshman, 2008; Mondary, 2008). No entanto, reconhecem-se várias limitações a esta abordagem, a começar pelo próprio processo de selecção e delimitação dos *corpora* utilizados para aceder à informação, cujo desenho e dimensão pode, na nossa opinião, ter um impacto significativo nos resultados do processo de acesso, extracção e representação do conhecimento.

Para além deste desafio, este tipo de abordagem enfrenta, na nossa opinião, outros problemas ligados, por exemplo, à avaliação dos dados extraídos e, conseqüentemente, à utilidade e possibilidade de reutilização da informação identificada como relevante, dado que não podemos afirmar que o conjunto de candidatos a termo que podemos encontrar e extrair, de modo automático ou semiautomático, de um texto de especialidade, seja o reflexo de uma estrutura organizada de conceitos.

Na perspectiva de Costa (2010: 03), o trabalho de representação do conhecimento de um domínio partindo do texto é altamente questionável. Num texto de especialidade são denominados conceitos e são estabelecidas relações lexicais que denotam parcialmente as relações conceptuais. No entanto, o texto é, ele mesmo, um sistema complexo que, na sua opinião, não se sobrepõe à complexidade do conhecimento representado. Como deixa claro recorrendo às palavras de (Rastier, 2001: 21) não é um:

suite de schémas cognitifs (propositions mentales, modèles mentaux, scripts, plans, etc.). Sa lecture suscite certes des corrélats mentaux, mais sa structure ne consiste pas en de tels corrélats. Par rapport à la psychologie, il apparaît non pas comme un ensemble de

représentations, mais comme un ensemble structuré de contraintes sur la formation des représentations.

Como afirma, os especialistas, no que respeita ao léxico, empregam, na transmissão do conhecimento e em situações comunicativas de especialidade, unidades que, de um ponto de vista linguístico, possuem estatutos diferentes, assumindo, em paralelo, funções diferenciadas: denominações, colocações terminológicas, combinatórias lexicais, palavras da língua corrente. Segundo Costa (2010: 03), todas estas unidades são indispensáveis para a construção do discurso, o que não quer dizer que o sejam para a representação formal do conhecimento.

Esta perspectiva vai de encontro ao ponto de vista expresso por Roche (2006: 01) de que o texto não é uma *tradução directa* para uma língua natural de um domínio de conhecimento. Escrever um documento cai na esfera do que Roche define como “língua em acção”, o que implica que o texto está mais próximo de uma:

‘linguistic picture’ in a given context at a given time of a state of domain knowledge than a ‘specification’ of the invariant knowledge of this domain. Text analysis is the concern of textual linguistics of which one of the principles is the incompleteness of text.

Neste sentido, compreender o texto e, em seguida, o conhecimento que transmite, requer conhecimento extralinguístico que, por definição, não se encontra no corpus. Este ponto de vista está patente nas palavras de Lerat (2008: 74) quando afirma que, se admitirmos que os dados correspondem a representações simbólicas de factos, que as informações correspondem a dados inseridos em contexto e que o conhecimento corresponde a informações interligadas entre si, podemos afirmar que os textos de especialidade

véhiculent des connaissances spécialisées, où l’analyse des contextes conduit à des informations, dont la décontextualisation aboutit à des données on voit clairement pourquoi l’expertise du domaine ou du métier est nécessaire : sous les mots il y a des informations, sous les informations des connaissances.

Como refere Costa, uma vez que a lógica do discurso se procura apropriar da lógica subjacente a uma conceptualização, o texto é frequentemente entendido não como uma representação do conhecimento, mas como o próprio conhecimento. No entanto, o conhecimento, entendido como “*l’acte de la pensée qui saisit un objet par les sens ou non*

avec implication plus ou moins grande du sujet de la connaissance” está centrado sobre o sujeito e não sobre o texto (Calberg-Challot *et al.*, 2009: 35); o texto é a representação de uma percepção mental que contém uma representação não formal, codificada linguisticamente, de um conhecimento.

Por outro lado, como salienta Costa (2010: 01), a ambiguidade e a polissemia são características intrínsecas ao discurso em língua natural que os sistemas formais procuram reduzir ou anular. No entanto, os especialistas não podem evitar nenhuma daquelas, uma vez que o discurso continua a ser o meio privilegiado de comunicação para transmitir e partilhar conhecimento com os membros da sua comunidade profissional.

De facto, o saber que o indivíduo detém sobre uma ciência consiste, na essência, no conhecimento e na capacidade de criação de discursos e de textos que poderão servir de referência a outros ou a futuros membros da comunidade científica, sendo através do discurso, oral e escrito, que o especialista transmite a sua visão do mundo e que representa *“la conceptualisation langagière des choses qu’il veut partager”*, permitindo o discurso que se situe face à comunidade científica a que pertence (Costa, 2010: 02).

Assim, e embora consideremos que o recurso a uma perspectiva linguística, de base semasiológica, não seja suficiente ou adequada para formalizar conhecimento num sistema computacional, o trabalho terminológico realizado com base nesta abordagem pode contribuir, na nossa opinião com informação sobre os conceitos e as suas denominações, abarcando, assim, como afirma Santos (2010: 89), *“o linguístico e o extralinguístico, uma vez que o significado dos termos contém também uma significação ligada ao acto comunicativo e cultural da sua produção, ligada, portanto, a uma parte conceptual e extralinguística”*⁴¹.

É, na nossa opinião, uma abordagem que deve estar presente no desenrolar do trabalho terminológico, sobretudo se considerarmos os mais recentes avanços tecnológicos que têm permitido o seu aperfeiçoamento e a melhoria de alguns dos resultados obtidos. Importa, no entanto, a nosso ver, ter presente e reconhecer as limitações de uma abordagem exclusivamente de base linguística, uma vez que esse reconhecimento torna

⁴¹ Para a autora, uma perspectiva puramente conceptual, de base onomasiológica, centrada somente no conceito e no extralinguístico, corre o risco de cortar a ligação dos conceitos às conceptualizações, uma vez que uma conceptualização, do ponto de vista cognitivo, está ligada ao locutor, à sua cultura e à língua natural em que se expressa, ou seja, ao discurso, objecto privilegiado da análise linguística, de base semasiológica. (Santos, 2010: 88)

mais claro o modo como se pode otimizar e potenciar o uso dos textos como apoio ao processo de conceptualização.

Os textos podem revelar-se como um elemento de grande utilidade, se manipulados com cuidado e, como deixámos claro, se manipulados com a noção clara de que um texto de especialidade pode não corresponder ou conduzir à obtenção de elementos que sustentem a construção de uma estrutura organizada de conhecimento, sobretudo no caso da construção de uma ontologia.

1.5.2. Perspectiva conceptual

A perspectiva conceptual centra-se no conceito. O recurso a uma perspectiva conceptual implica que a análise do conhecimento do domínio se realize a partir do conceito, sendo a primazia do conceito sustentada por elementos conceptuais que se reportam essencialmente ao conhecimento transmitido pelo especialista de domínio.

Na perspectiva de Roche (2008) a língua natural não representa um veículo adequado para especificar formalmente uma conceptualização⁴². Esse não é o seu objectivo, tal como o objectivo de escrever um texto não é o de definir uma ontologia. Como refere, as coisas não são tal e qual como são ditas: « *dire n'est pas concevoir* » (Roche, 2008b). Para o autor, a conceptualização de um domínio não pode ser extraída de uma única análise linguística de textos, dado que frequentemente o conhecimento não se encontra escrito em documentos científicos e técnicos (por exemplo, conhecimento tácito), o que torna a presença dos peritos e a sua colaboração indispensável ao trabalho terminológico e à organização do conhecimento, opinião com a qual concordamos.

Roche defende, assim, que um modelo conceptual construído a partir dos textos não corresponde a uma ontologia. Uma tal conceptualização, assente numa perspectiva linguística, seria, do seu ponto de vista, dependente do *corpus* e incapaz de oferecer as

⁴² A terminologia é, nas palavras de Roche (2005: 50), uma língua “*de l'intellection qui étudie les rapports entre les concepts et les choses*”, o que o leva a defender como ponto de partida para a organização do conhecimento de um domínio, uma abordagem terminológica centrada na conceptualização do conhecimento. Nas palavras de Roche, tal como a língua é considerada um sistema, também uma terminologia é um sistema, neste caso de termos que reflectem uma conceptualização.

principais propriedades expectáveis numa ontologia, como a reutilização. Para além disso, afirma que uma ontologia extraída a partir de textos não corresponde às ontologias definidas por especialistas com recurso a uma linguagem formal.

Esta não correspondência resulta, no seu ponto de vista, do facto de uma ontologia ser o resultado de uma conceptualização extralinguística, enquanto que o conhecimento extraído dos textos:

is the concern of textual linguistics. Incompleteness of text and using rhetorical figures, like synecdoche, deeply modify the perception of the conceptualization we may have. It means that ontological knowledge, which is necessary for text understanding, is not in general embedded into documents. (Roche, 2007: 47)

Por outro lado, e tendo em conta que as aplicações relacionadas com o tratamento da informação necessitam de uma representação computacional dos conceitos, considera que a terminologia não consegue oferecer os meios necessários para a sua operacionalização. De acordo com o seu ponto de vista se uma conceptualização “*se dit bien en langue, elle se définit dans un système formel guidée par des principes épistémologiques où l’ontologie occupe une place prépondérante*”⁴³ (Roche, 2008: 70). A utilização de uma linguagem formal, cuja sintaxe e semântica sejam claramente definidas torna-se, assim, essencial para eliminar as ambiguidades⁴⁴.

Esta posição está próxima da de Grenon (2008: 76) quando afirma que a língua natural fornece pistas para a construção de uma ontologia, mas não deve ser o único critério de correcção e de validação do resultado final desta construção, posição na qual nos revemos. Tal como nos lembra Grenon, se tudo na língua natural correspondesse a algum aspecto da realidade, então não haveria, à partida, necessidade de proceder à representação ontológica ou à modelação conceptual.

⁴³ Roche chama a atenção para o facto de a lógica (Baader *et al.*, 2003) e das linguagens formais utilizadas pela IA (Sowa 2000), (Brachman *et al.*, 1985) serem antes de tudo, sistemas de representação cuja finalidade não é a de compreender o mundo, mas de o descrever de “*manière formelle pour les premiers, à des fins de calcul par un ordinateur pour les seconds*”.

⁴⁴ Roche defende que a sociedade numérica, ao reclamar a operacionalização das terminologias para proceder à gestão da informação, reactualizou a terminologia Wüsteriana. A dimensão conceptual reencontrou, na sua perspectiva, um lugar preponderante que tinha perdido em favor da lexicografia de especialidade. Dá, como exemplo desta afirmação, a emergência da noção de ontologia no seio da terminologia.

Como afirma, em consonância com Roche, um modelo conceptual construído directamente a partir de uma estrutura lexical poderá não ser reutilizável, uma vez que é demasiado dependente do *corpus*, e, por outro lado, incorrecto, se se considerar uma ontologia como *a non-contingent knowledge*. Para Grenon (2008: 77), se a língua natural tiver a capacidade de impor “*an ontological inquiry unimpeded, the ontologist risks making the radical claim that an ontology must stand in a one-to-one correspondence with the elements of natural language (or as close thereto as possible)*”. Tal conduziria, no limite, a que obtivéssemos uma relação do tipo uma palavra - um conceito (*‘one word, one concept’*), resultando daqui um sistema demasiadamente rico em conceitos e distante de uma ontologia formal.

A perspectiva conceptual, adoptada por algumas abordagens da inteligência artificial e da Terminologia encara, assim, uma ontologia como uma construção onde o especialista e o engenheiro do conhecimento ocupam o papel central na conceptualização, resultando deste processo uma representação da realidade diferente da que se poderia eventualmente extrair a partir de elementos linguísticos presentes no texto. O recurso a uma perspectiva conceptual implica, assim, reconhecer que não é possível construir ontologias de modo directo a partir de textos ou de *corpora*.

De facto, e dadas as limitações inerentes ao uso de textos, que apontamos antes, importa reconhecer que no texto não existe conhecimento, mas antes elementos linguísticos que fornecem pistas para a organização conceptual de um domínio, o que inviabiliza, por exemplo, o uso dos textos como suporte à construção automática ou directa de ontologias, onde podem mesmo constituir-se como elementos que geram algum ruído. O recurso ao texto encerra, efectivamente, dificuldades, criando obstáculos associados à utilização de *corpora* para a realização de um trabalho terminológico de representação do conhecimento de um domínio, até porque, não raramente, “*on trouve dans les textes ce que l’on veut y voir et non pas ce qui y est*” (Costa 2010: 03).

No entanto, e como afirma Santos (2010: 101), na abordagem conceptual, o texto pode ser prescindível, mas não a terminologia. A terminologia, partindo deste ponto de vista, está para além do texto e não se deve prender ao discurso verbal sob pena de não poder ser partilhada e reutilizada.

O recurso a uma perspectiva conceptual implica que a valência do conhecimento se centre no conceito e não no termo. As relações entre os conceitos reflectem a organização do conhecimento de domínio, cuja rede é elaborada com recurso prioritário ao especialista que trabalha em conjunto com o terminólogo ou com o engenheiro de ontologias, sendo a especificação da conceptualização, referida por Gruber, feita através de uma linguagem, e não da língua. Santos (2010: 94) entende, a este propósito que *“o que tendencialmente se verifica é que a linguagem de representação descreve, mas não define, o conceito”*.

A importância da conceptualização de um domínio é vista por Roche (2008: 70) como fundamental para a terminologia, uma vez que *“il ne peut y avoir de terminologie sans connaissances”*. Na nossa perspectiva, o inverso é também verdade, i.e., a terminologia é um elemento essencial de apoio a uma conceptualização efectiva de um domínio de especialidade e essencial ao processo de conceptualização que permite e promove o acesso ao conhecimento desse domínio, tal como defendemos antes.

Entendemos também, como deixámos claro antes, que a participação do especialista no processo de conceptualização - uma preocupação central ao uso de uma abordagem conceptual e essencial para o seu sucesso - não é isenta de riscos, pelo que se torna essencial fornecer os meios adequados e acompanhar o especialista do domínio, de modo a que este forneça o conhecimento necessário à construção do sistema conceptual sem necessariamente compreender as propriedades formais ou computacionais ou a linguagem formal de representação do conhecimento subjacente à sua formalização.

Na nossa opinião, as duas perspectivas, a linguística e a conceptual, não ocupam planos antagónicos ou mutuamente exclusivos. Consideramos mesmo que uma abordagem mista, que resulte do uso complementar das duas perspectivas, de modo flexível e em adequação às necessidades e propósitos de análise do conhecimento de um domínio específico, pode conduzir a resultados mais profícuos na representação desse mesmo conhecimento. Tal como afirmava Costa já em 2005:

Si d'un côté, il y a symbiose, de l'autre la distinction théorique est indispensable pour comprendre la différence entre la perspective linguistique, plutôt sémasiologique et la perspective conceptuelle, plutôt onomasiologique, qui ne s'excluent pas mutuellement, mais si complètement, ceci étant finalement de la compétence du terminologue. Cette opposition presque dichotomique nous semble aujourd'hui bien réductrice.

Por outro lado, ainda que a formalização das representações obrigue a que as especificações das conceptualizações sejam feitas através de uma linguagem, não da língua, resultando daí um conjunto de restrições, e que não exista uma relação de um para um entre texto e ontologia, acreditamos, tal como (Santos, 2010b: 02), que a construção de ontologias *“não pode nem deve subtrair-se à questão linguística, sendo o contrário igualmente verdadeiro: o lexical não deve substituir o conceptual. Assim, na construção de ontologias não existe exclusividade, mas complementaridade, entre as abordagens que partem do linguístico e do extralinguístico”*.

Como afirmam Costa e Silva (2008), sendo verdade que a organização do conhecimento se encontra ao nível extralinguístico, não é menos verdade que:

*c’est à travers l’acte de la parole, c’est-à-dire le discours, que nous pouvons accéder à la représentation de l’organisation des connaissances. La parole étant un moyen privilégié de représentation du monde et l’endroit de construction du sens; mais elle n’est pas le monde en soi.*⁴⁵

Já no que se refere à especificação de uma ontologia em mais do que uma língua natural, acreditamos que o recurso às duas perspectivas pode resultar numa mais valia no desenvolvimento de uma abordagem interlinguística, ainda que a perspectiva conceptual, decorrente da abordagem onomasiológica, deva assumir a primazia.

Acreditamos que, numa abordagem interlinguística, em que se parte do princípio que a representação conceptual é transcendente às diferentes línguas, o processo de escolha dos termos não pode partir das formas, mas dos conceitos, não podendo, conseqüentemente, ser uma tradução ou uma localização de nomes, ou uma simples questão de equivalência das denominações, mas antes uma busca pela denominações naturais que em cada língua correspondam a um conceito de especialidade, tal como expresso pelos especialistas em situações de comunicação profissional, o que reforça a necessidade da sua presença e do seu papel de validação.

⁴⁵ Os signos usados para nos referirmos à realidade não são, de facto, a realidade em si mesma, mas apenas a sua representação mental organizada em conceitos (Cabré, 1999: 40). No entanto, estas duas realidades – o mundo e a sua representação discursiva – formam para Costa e Silva (2008) o que descrevem como uma associação estável e indissociável, revendo-se as autoras nas palavras de Aubenque (1962:128), quando diz *“C’est parce que les choses ont une essence que les mots ont un sens”*.

1.6. Representação do conhecimento em contextos multilingues

Comparison of languages relies crucially on those concepts which are essentially the same across human minds, cultures, and languages, and which therefore can be activated through the use of any human language. These instances of mental universals join other less common concepts to constitute a complex structure in our minds, a network of cross-connected conceptualizations of the phenomena that make up our world. (Schalley et al., 2007: 03)

O desenvolvimento de organizações virtuais e a sua passagem à prática requer a criação de processos e aplicações que apoiem a comunicação e a compreensão entre as pessoas. Daí que subsistam problemas, sobretudo no processo de colaboração e partilha do conhecimento, onde a língua e a cultura desempenham um papel fundamental. O conhecimento a nível internacional transmite-se através da língua, estando a dimensão conceptual deste processo presente na conceptualização e na transferência de conhecimentos.

Um conjunto importante e complexo de factores interligam-se para garantir o sucesso de um espaço de trabalho multilingue e multicultural. De facto, o trabalho em equipas multilingues está cada vez mais disseminado e a interacção comunicativa torna-se essencial, quer para a dinâmica da equipa e para a prossecução dos seus objectivos quer nos aspectos relacionados com a partilha, representação e disseminação do conhecimento.

Como descreve Nakata (2007: 199) as redes interorganizacionais envolvem frequentemente uma multiplicidade de equipas, oriundas de organizações diversas, que possuem culturas, visões do mundo e práticas profissionais próprias, bem como conceptualizações e compreensões nem sempre coincidentes do domínio de especialidade, sendo importante criar espaços e meios que permitam a intersecção entre os diferentes interlocutores, as suas histórias, experiências, línguas, aspirações e propósitos e que promovam a partilha de conhecimento, o estabelecimento de redes e a aprendizagem colaborativa.

No entanto, a interacção do ser humano com o mundo é dependente da sua compreensão e conceptualizações dos objectos do mundo. As pessoas, quando inseridas em diferentes culturas e em diferentes países têm a tendência para conceptualizar as coisas de

modo diferente, e, por implicação, a classifica-las de modo diferente, diferença que, não raramente, transparece entre os interlocutores de uma determinada comunidade.

Para que seja possível partilhar e reutilizar conhecimento, torna-se necessário alcançar, como descrevemos, um elevado grau de consenso, sistematização e coerência, que dependem, igualmente, da aplicação e da utilização dos dados e da informação e de pressupostos temporais e operacionais exigentes. Neste contexto, o conhecimento, e a sua representação, adquirem quase sempre contornos funcionais.

Para Montiel-Ponsoda (2011: 01), pode-se argumentar que o objectivo de uma ontologia é o de representar, “*a priori nature of the world, avoiding any type of arbitrariness or partial view imposed by languages*”. De facto, o conhecimento representado numa ontologia deve corresponder a uma representação livre de propriedades acidentais e baseado nas características essenciais dos elementos. No entanto, e como lembra esta, o que se inclui numa ontologia captura, não raramente, o que vai de encontro aos interesses de uma determinada comunidade ou que melhor condiz com os propósitos de determinada aplicação.

Na sua perspectiva, com a qual concordamos, as ontologias formalizam uma determinada visão do mundo tal como é entendida por uma comunidade de utilizadores, tendo em conta um conjunto de propósitos, sendo a língua um veículo que contribui para esta construção, uma vez que reflecte o modo como uma comunidade de utilizadores compreende a realidade.

Assumindo que as ontologias representam conceitos e não o modo como estes conceitos são expressos em palavras numa língua natural, Weigand *et al.* (2008) consideram natural que se assuma que uma ontologia é independente da língua em que é criada. Apontam, contudo, problemas a esta assunção, a começar pelo ponto de vista filosófico, segundo o qual não é possível ao ser humano sair do seu contexto linguístico, dado que os conceitos ganham forma na comunicação que se gera entre membros de uma comunidade linguística, pelo que, acrescentam, não é natural que se *desligue* a ontologia da língua.

Na nossa perspectiva, e tendo em conta que um conceito é uma *unidade de conhecimento* criada por uma combinação única de características (ISO 1087-1: 2000,

assumimos que os conceitos que não estão, necessariamente, ligados a uma língua em particular. Como destaca Roche (2009: 54):

Le concept structure la réalité de manière stable, indépendamment de la langue – il y a ici bi-univocité entre le concept et sa dénomination. Il ne doit pas être confondu avec le mot – un terme s’emploie comme un mot en texte – dont le signifié découpe de manière contingente une réalité construite en discours.

Os conceitos são, no entanto, influenciados pelo *background* social e cultural que frequentemente conduz a diferentes categorizações (ISO 1087-1:2000), facto que devemos ter presente no desenrolar do processo de conceptualização, sobretudo quando este acontece no seio de uma comunidade multilingue.

Não devemos, contudo, abordar o conceito, enquanto entidade independente de uma qualquer língua, a partir de um ponto de vista que reflecta as distinções que cada língua particular possa fazer no seu vocabulário ou na sua análise da realidade. Devemos antes, como aponta Antia (2000: 84), seguir uma abordagem onomasiológica, abordagem que, como descreve, corresponde à totalidade das estratégias através das quais uma análise terminológica revela:

the ontology of knowledge in specialised knowledge areas in a manner that is independent of a given language, but lets the analyst identify any biases of a given subsection of the knowledge community – whose self definition may very well coincide significantly with a given language space. (...) Onomasiological strategies are also useful in those cases where a given specialised knowledge is the exclusive preserve of a given language space.

No âmbito da especificação multilingue de uma ontologia tanto a representação como a transferência do conhecimento estão restringidas por factores de conveniência e consenso, factores que são extensíveis à construção da representação ontológica do conhecimento em mais do que uma língua natural, onde desempenham um papel decisivo. Neste processo intervêm, normalmente, terminólogos e especialistas, que procuram estabelecer, através de consenso, uma unidade no que se refere à “*fijación de unas equivalências interlingüísticas no sempre fáciles de establecer* (Cabré, 2004: 98)”⁴⁶.

⁴⁶ Como afirma Briu (2012), “*En étudiant, on découvre, avec le temps, que le sens des mots, des signes, des termes et des textes est plus complexe qu’on ne croyait, que les langues sont des systèmes souples, très puissants et que la traduction, qui met en rapport au moins deux langues, est une opération naturelle, possible, mais hyper-complexe et risquée, un défi ou une « négociation ».*”

Neste sentido, destaca-se o papel da terminologia neste processo. Tal como afirma Budin (2005:106), as terminologias capturam de forma mais profunda a relatividade intercultural e muitas das assimetrias existentes entre *“culture-specific conceptualizations and their lexicalizations on the cross-lingual level”*. Esta característica leva-nos a considerar o recurso à terminologia como um elemento essencial no processo de representação do conhecimento em contextos multilingues ou de construção de uma ontologia de domínio que represente conhecimento em mais do que uma língua natural.

Como afirma Bessé (1997: 73), o factor multilingue e multicultural é recorrente em terminologia. Tal não conduz, no entanto, a que os terminólogos, e eventualmente, outros actores envolvidos no processo, traduzam os termos ou as definições. O processo passa antes por identificar o conceito - que pode ser diferente de uma língua para a outra, ou não existir na língua de chegada -, descrevê-lo de acordo com o sistema de conceitos da cultura de chegada e, finalmente, caso haja necessidade, encontrar uma designação, ou seja, um termo adequado à língua natural em questão.

De facto, dadas as especificidades dos sistemas linguísticos, nem sempre é possível estabelecer uma relação de equivalência precisa entre termos de línguas diferentes. À falta de um equivalente adequado na língua alvo, a solução de tradução passa frequentemente pelo recurso a um empréstimo ou estrangeirismo a uma perífrase ou a uma paráfrase. Neste contexto, a terminologia desempenha um papel primordial no respeito pela concisão, precisão e adequação da comunicação especializada, devendo os conceitos, sempre que possível, serem nomeados

de telle sorte que les noms reflètent la structure notionnelle. L'organisation du système notionnel et les principes épistémologiques sur lesquels repose la terminologie ont donc un impact jusqu'à la denomination des concepts. (Roche 2005: 54)⁴⁷

⁴⁷Como especifica Van Campenhoudt (2006: 06), em resposta à questão *L'équivalence: concept ou monosemie?* no seu artigo *“Que nous reste-t-il d'Eugen Wüster?”*, Wüster demonstrava já preocupações com a tradução, sendo que a sua abordagem, apesar de criticável, dado o seu carácter prescritivo, oferecia um modo de lidar com os problemas de equivalência. Nas palavras de Van Campenhoudt (2006: 06), *“Mettre en avant la définition et le classement typologique conduit, en effect, à garantir que les termes désignent bien les mêmes choses.”* De notar, por outro lado, que a abordagem normalizadora da terminologia, proposta por Wüster, surge num momento histórico em que se tornava essencial a clareza na transferência do conhecimento e havia grande necessidade de fixar as designações dos novos conceitos, com o intuito de evitar dificuldades de comunicação.

Esta sistematização das denominações é relevante, tal como descreve Costa (2010), dada a correlação directa entre os sistemas conceptuais e a qualidade dos sistemas denominativos, permitindo os primeiros *“une meilleure production de textes de spécialités, donc une communication plus rigoureuse entre spécialistes, donc de meilleures traductions”*.

Tendo, assim, por base uma perspectiva conceptual para o desenvolvimento da especificação multilingue da ontologia, entendemos que os termos podem ser vistos como unidades que cumprem uma dupla função, a representativa e a comunicativa. Na sua função de representação denominam uma realidade construída com base num processo de negociação e consenso, o que acontece quando se constrói uma ontologia. Por outro lado, e do ponto de vista da transferência, permitem denominar interlinguisticamente a dita realidade construída a partir de um consenso sobre as formas de equivalência.

Apesar da importância da análise interlinguística para a correcta construção de uma representação do conhecimento do domínio, acreditamos, como explicitámos já, que enquanto elemento de representação, o sistema conceptual é extralinguístico, o que não significa, no entanto, que tenha que ser forçosamente interlinguístico. De facto, um sistema conceptual que represente o conhecimento de um domínio nem sempre é comum às várias línguas, sendo normalmente concebido como independente e não isomórfico face a qualquer língua, pelo que não tem que ser necessariamente interlinguístico.

Admitimos, no entanto, tal como Roche, que a existência de um sistema conceptual interlinguístico se torna viável, quando é construído no seio de uma comunidade *“dont les langues d’usage partagent la même conceptualisation du monde”* (Roche, 2005: 54). Para o autor, a compreensão (tradução) entre diferentes comunidades poderia ser vista da perspectiva de uma *“« mise en correspondance » de systèmes notionnels différents”*.

Acreditamos, por outro lado, que a selecção da terminologia adequada, tendo por base conceitos bem definidos e o uso correcto e consistente dos termos representa uma pré-condição para o desenvolvimento efectivo dos sistemas de gestão de informação baseados em ontologias, o que demonstra a necessidade da existência de um processo de gestão da terminologia ao longo de todo o processo de conceptualização e de localização do conhecimento.

Para que este processo de construção de uma ontologia ocorra dentro de uma comunidade multilingue e para que se torne possível a partilha e a análise do conhecimento do domínio, o processo de conceptualização da ontologia socorre-se de uma língua pivô, geralmente a língua inglesa. O uso deste tipo de estratégia permite criar um artefacto que procura – dado o papel de língua de trabalho que assume a língua inglesa – uma quase *desterritorialização* do artefacto e, por conseguinte, do conhecimento. O uso de uma língua pivô procura contribuir para que o resultado da construção da ontologia, enquanto representação do conhecimento do domínio, assuma as necessárias características de consenso entre as várias culturas presentes na comunidade.

O resultado da conceptualização usando uma língua pivô não deve, no entanto, ser perspectivado como um resultado intermédio – é um artefacto com valor próprio, negociado e validado por uma comunidade, que ocupa o seu lugar na construção do conhecimento, sobretudo quando estamos perante comunidades que há muito adoptaram a língua inglesa como elemento não só de comunicação e de disseminação do conhecimento, mas também de construção do próprio saber.

Este artefacto corresponde a uma estrutura conceptual obtida por consenso, estável, que representa o conjunto de conhecimento sobre o domínio e subdomínios em análise sendo, em muitos casos, utilizada como estrutura nuclear para permitir a interligação entre as ontologias representadas nas diversas línguas naturais necessárias à representação do conhecimento numa comunidade multilingue.

Esta perspectiva, que aponta para uma metodologia de análise interlinguística, assente no trabalho interdisciplinar, desenvolvido por terminólogos em conjunto com especialistas do domínio, num espaço de partilha multilingue e multicultural, corresponde, a nosso ver, à abordagem a seguir na conceptualização do conhecimento do domínio num contexto multilingue.

Acreditamos que esta abordagem é importante para o processo de conceptualização e para a representação multilingue do conhecimento de um domínio, dado que constitui um meio para a recolha dos aspectos transculturais e das assimetrias entre as conceptualizações resultantes de cada cultura e das suas lexicalizações ao nível interlinguístico.

CAPÍTULO II

2. Localização

A indústria da localização, que emergiu na década de oitenta do século XX, sofreu um avanço intenso na década seguinte, fomentado pelo mercado dos computadores pessoais e pela necessidade de localizar o grande número de *softwares* que entretanto surgiram. A principal diferença apontada nessa altura entre a tradução e a localização era a de que na localização as cadeias de caracteres (*strings*) e as cadeias de texto (*text strings*) traduzidas eram embutidas no *software*, sendo o processo de localização conjugado com um conjunto de procedimentos novos, como o teste ao produto localizado, o que resultava numa combinação entre tradução e engenharia de *software* e na conseqüente exigência de novas competências aos tradutores.

Desde então, o processo de localização tem vindo a amadurecer e a abranger novas áreas de utilização, graças sobretudo ao avanço técnico e tecnológico promovido e procurado pela indústria da língua, à medida que as práticas de localização vão evoluindo em resposta às necessidades de globalização de uma multiplicidade de conteúdos digitais. No entanto, esta dinâmica não tem encontrado grande paralelismo na análise e estudo quer do conceito quer do processo de localização junto da academia, apesar da sua importância e das perspectivas que abre sobre novas dimensões e caminhos para o futuro do fenómeno da tradução, enquanto processo ligado à gestão e disseminação de conteúdos digitais e à transferência de conhecimento a nível global.

Com a evolução da indústria do *software*, foram-se, no entanto, definindo um conjunto de conceitos nucleares, reunidos em torno de um novo acrónimo – GILT - que incorpora a Globalização, a Internacionalização, a Localização e a Tradução, acrónimo que reflecte a complexidade inerente à preparação de um produto ou conteúdo para o mercado global. Segundo O'Hagan (2004: 02) GILT, para além de destacar os processos específicos necessários ao tratamento de conteúdos electrónicos, coloca a transferência dos conteúdos digitais provenientes das diversas línguas no espectro mais abrangente da globalização. Interessa, por isso, explicitar melhor os conceitos abrangidos pelo acrónimo, as suas interligações e o modo como se reflectem nos processos de localização e de tradução e como influenciam a transferência de informação e conhecimento entre culturas.

A IBM define, no seu glossário sobre internacionalização, o conceito de *globalização* como sendo o processo de desenvolvimento, produção e marketing de produtos (de *software*) a serem distribuídos em todo o mundo⁴⁸. Este conceito engloba a internacionalização (criar um produto que possa ser usado livre de barreiras linguísticas e culturais) e a localização (tradução e adaptação de um produto a um local específico)⁴⁹. Esta definição aclara a interligação e dependência existente entre os conceitos englobados pelo acrónimo GILT, que podem ser representadas do seguinte modo:



Fig. 2 - Conceitos GILT (Adaptado de Hoft:1995)

A globalização pode, assim, ser descrita como um processo de criação de um produto passível de ser usado com sucesso em diversos contextos culturais sem sofrer modificações. Autores como Esselink (2000) ligam ainda esta perspectiva à dos processos e estratégias de negócio associados à criação de um produto global.

A internacionalização surge associada ao processo de globalização e é descrita por Hoft (1995) como o processo de reengenharia de um produto de informação de modo a que possa ser facilmente localizado para exportação, sendo um produto internacionalizado composto por duas componentes: informação nuclear ao produto e variáveis internacionais. Esselink (2000: 02), por seu lado, descreve a internacionalização como o processo de generalização de um produto, desenvolvido para que este possa lidar com uma multiplicidade de línguas e de convenções culturais sem haver necessidade de o redesenhar. Corresponde, assim, a um processo de desenvolvimento de um produto de modo a que este não seja concebido especificamente para um local, i.e., que possa suportar diferentes

⁴⁸ <http://www-03.ibm.com/systems/i/software/globalization/quickterms.html>

⁴⁹ <http://www.ibm.com/developerworks/library/glossaries/unicode.html>

línguas e convenções, independentemente da sua origem, sem que, para tal, necessite de modificar qualquer componente do seu *design* inicial.

Este processo antecede o de localização. Na origem do conceito de localização está o termo inglês *locale*. Este termo estendeu o seu significado a partir da acepção da língua geral do substantivo *locale*. Quando se fala em *locale*, no contexto de um sistema operativo ou de uma aplicação informática, faz-se referência a uma área geopolítica e a uma língua determinada, bem como às restantes particularidades que devem ser consideradas – no caso de uma aplicação, o modo de expressar as horas e as datas, as letras do seu alfabeto, etc. (Pagans, 2000).

Autores como Texin (2003) e Guidère (2004) vêem um *locale* como um grupo de preferências culturais. Um *locale* é, segundo Texin (2003: 26), descrito por um conjunto de características:

- *world* – representing the cultural preferences of the users in the whole world;
- *distinctive* – permitting differentiation of the various groups of users;
- *correlative* – connecting a culture to a locale;
- *determinative* – correctly describing the types of preference linked to each locale;
- *specific* – defining accurately the specific preferences of all locales.

A localização, elemento nuclear e essencial, no nosso ponto de vista, à globalização e à localização, é definida como um processo de adaptação linguística e cultural de informação digital para mercados locais (Esselink, 2000; Dunne, 2006). O facto de estes autores se referirem a informação e a conteúdos digitais afasta-os da perspectiva mais restritiva de Hoft (1995), segundo a qual a localização se dedica ao processamento de *produtos de informação*, enquadrando estes, quase só, as aplicações típicas de *software*.

A localização pode, então, ser caracterizada como um processo de adaptação (1) de carácter linguístico, com recurso à tradução, à terminologia e à edição, e (2) de carácter não-linguístico, pelos processo de programação e ergonomia envolvidos, tendo como objectivo último o desenvolvimento de um produto aceite por uma determinada comunidade linguística e cultural.

Este conceito tem vindo a evoluir à medida que o próprio processo e a indústria evoluem e Schäler (2007: 157) aponta agora uma nova dimensão, na qual nos revemos. Para

este autor, a localização representa não só a adaptação linguística e cultural de conteúdos digitais aos requisitos de um mercado internacional, mas também o fornecimento de serviços e tecnologias para a gestão do multilinguismo através do fluxo de informação digital a nível global.

A localização necessita, assim, de ter em consideração a comunicação intercultural e factores como o estatuto de uma língua, a política linguística, as preferências culturais, os usos estabelecidos e, até mesmo, as necessidades dos utilizadores, centrais ao processo de aceitação, uso e sucesso de qualquer informação ou produto.

2.1. Localização e Tradução

A evolução das teorias da tradução procurou consolidar a noção de um original inalterável bem como a de um autor como sujeito proprietário do texto por ele escrito (Bassnet e Trivedi, 1999: 02). Perante este texto, visto como um produto acabado, a tarefa do tradutor consiste, essencialmente, em produzir um texto traduzido que represente e seja ‘fiel’ ao texto original.

Embora a noção de fidelidade tenha quase tantas definições como teóricos da tradução e apesar de as correntes mais recentes como o desconstrutivismo colocarem em causa a noção de ‘original’ e de ‘autor’ (Davis, 2001), o facto é que o tradutor está limitado na sua estratégia tradutiva e nas suas escolhas lexicais, sintácticas e estilísticas, por esta necessidade de fidelidade ao texto de partida.

Durante séculos prevaleceram diferentes teorias prescritivas de tradução, cuja origem pode remontar ao poeta latino Horácio (Lefevere, 1992:15) e que encontramos ainda no teórico Peter Newmark (1988, 2007), e perpetuadas pela indústria editorial e pelas publicações especializadas, que se encarregaram de salientar a necessidade de preservar, tanto quanto possível, forma e conteúdo para conseguir um “efeito equivalente” (Nida, 2000).

Com o surgimento da localização, a fidelidade ao original passou a assumir um plano quase secundário, sobretudo porque o original, no sentido tradicional, deixa de existir, isto é, o texto localizado já não representa um texto original acabado e preexistente (Pym,

2004a: 5). Esselink (2000: 2) aponta algumas razões para essa evolução e destaca como diferença chave o facto de a tradução, que chama de tradicional, ser, tipicamente, uma actividade realizada depois do documento fonte ter sido concluído, enquanto na localização, os projectos se desenvolvem em paralelo com o desenvolvimento do “*source product to enable simultaneous shipment of all language versions*”.

Para além de o original assumir características de um objecto inacabado, sujeito a modificações constantes (e por isso mais fluido), deixa de existir um autor identificável como sujeito individual: no projecto de localização o autor do texto e do produto é um sujeito normalmente corporativo; o mesmo que inicia o processo e que, em última instância, valida a versão final.

É, portanto, natural que durante o processo de transferência linguística, os condicionalismos do localizador e do tradutor provenham de fontes diferentes: no caso do tradutor, da autoridade do autor e do texto original; no caso do localizador, dos objectivos do iniciador da localização e das limitações ou orientações explícitas impostas por este.

A tradução e a localização surgem assim como duas actividades diferentes. Embora não seja, nem nunca tenha sido, *apenas uma questão de língua*, a tradução, seja ela técnica, científica ou literária, ocupa-se sobretudo do discurso como veículo do pensamento. Dito de outra forma, para além de operar a transferência linguística de um *conteúdo* ou *mensagem*, ou precisamente para poder levar a cabo essa transferência, a tradução deve dar conta das intenções, atitudes, pressupostos e alusões que se manifestam no texto original através de um leque quase inesgotável de recursos, que passam pelas escolhas lexicais, as estruturas sintácticas, as relações de coesão, etc. (Tallone, 2010: 156).

A localização é, por outro lado, um processo que excede o da tradução e no qual a tradução é uma das fases, possivelmente não a mais importante (Pym, 2006: 05), para o qual contribuem diferentes profissionais que trabalham em diversos processos, do *design* ao controlo de qualidade. O objectivo principal desta transformação consiste, já não em permitir o acesso a textos que de outra forma estariam vedados aos leitores que não falassem a língua do texto de partida, mas, como explicitámos antes, na adaptação de um produto digital, através da sua localização, de forma a poder ser distribuído e utilizado em mercados distintos.

Para além disso, o facto de não existir um texto de partida único e fixo, mas uma sucessão contínua de modificações e actualizações, implica que na localização não se verifique uma linearidade do texto de partida. Tal como observa Pym (2002: 17), os textos já não podem ser vistos como elementos que operam como um todo orgânico, que resultam da mente de um criador único e coerente, mas antes aparentados com fragmentos:

(...) commonly fragments comprising fragments of other texts, in a language drawing on several languages, issuing from a series of reworking authors. (...) A text is quite simply whatever unit is distributed as a unit. In the age of printed books, one might have thought of books; in the age of combinable paragraphs, the text will be a 'chunk'; and so on.

Por outro lado, determinar qual é o leitor-alvo do texto a localizar é um processo complexo dado que, em vez de milhares podem existir vários milhões de leitores potenciais. Torna-se necessário, assim, que o texto localizado escolha um mínimo denominador comum e mantenha os princípios de simplicidade, clareza e neutralidade que orientam a produção do texto de partida.

Porém, o texto localizado não tem apenas leitores: os conteúdos localizados, enquanto partes integrantes de uma aplicação ou sistema, têm também utilizadores. Esta distinção entre leitores e utilizadores implica uma diferença na forma como são avaliados os processos de tradução e localização. Em tradução, o texto de chegada, até nas versões mais livres ou experimentais, assume-se como uma representação do texto de partida e a tradução é reconhecida como tal, na medida em que é possível estabelecer uma correspondência entre os dois textos (Tallone, 2010: 161). A confiança que o tradutor obtém do leitor e a 'qualidade' da tradução surge, segundo a autora, muitas vezes ligada ao modo como é percebida a proximidade ou afastamento do texto de chegada relativamente ao texto original.

Em localização, uma vez que o texto de partida é, não raramente, apenas um segmento descontextualizado, a confiança que o localizador merece é, como refere Pym (2004: 175), medida de acordo com o efeito do texto por ele produzido:

(...) such language workers ultimately achieve trust not on the basis of where their words have come from, but on what can be done with the texts they produce.

Apesar desta autonomia relativa do texto localizado, as escolhas do localizador estão longe de ser livres, ainda que as condicionantes sejam diferentes daquelas que balizam o

trabalho do tradutor. Enquanto este deve ter em conta questões de estilo, de registo, de tom, de intenção, etc., o localizador deve dar uma maior prioridade às restrições impostas pela aplicação e às orientações fornecidas pelo iniciador/cliente, sendo que, no caso da localização de um artefacto semântico, como uma ontologia, deve ter em conta ainda os pressupostos ontológicos (*ontological commitment*) que presidiram à sua construção e orientam a sua especificação.

Finalmente, no caso da tradução, o ónus da responsabilidade recai exclusivamente sobre o tradutor que é identificado e reconhecido como o produtor do texto de chegada. Em contrapartida, o texto localizado, como resultado de um esforço colectivo, remete, pelo menos em parte, o localizador ao anonimato.

2.2. Terminologia e localização

A relação entre tradução e terminologia foi já alvo da análise de vários autores (Cabré, 2004b; Budin, 2005; Rogers, 2007), tanto do ponto de vista prático, como teórico, análise que nem sempre foi extensiva ao processo de localização. De facto, o estudo dos termos e o estudo da tradução, sobretudo a especializada, e o estudo da localização gozam de uma relação próxima. Os termos constituem, como reconhece Rogers (2007: 01), *“an important semantic contribution to lexically dense texts dealing with specialist subject matter, and hence to their translation”*.

Terminologia, tradução e localização mantêm, enquanto disciplinas, uma relação de proximidade que se pode observar em diversos aspectos, como explicita Cabré (2004b: 89), que destaca a sua origem prática, a sua passagem relativamente recente a disciplina, a interdisciplinaridade constitutiva e a ligação imbrincada enquanto matérias relacionadas com a informação e a comunicação. Possuem, por outro lado, um carácter transdisciplinar, dado que transcendem um campo de saber, para estar presentes em todos os campos do conhecimento. A autora reconhece, no entanto, diferenças, em dois aspectos:

- a) O carácter *final* da tradução – que se estende, na nossa perspectiva à localização -, por contraste com o carácter *pré-final* da terminologia⁵⁰,
- b) A sua necessidade assimétrica específica.

Tanto a terminologia como a localização (e a tradução) são campos interdisciplinares e possuem na sua base elementos dos campos das ciências cognitivas, das ciências da linguagem, ciências da comunicação, da epistemologia e da lógica. A presença coincidente destes campos disciplinares resulta do facto de que o objecto das ciências dedicadas à comunicação e à informação se sintetizar a partir da descrição e explicação de três elementos:

- a) As categorias do conhecimento,
- b) As unidades expressivas que permitem expressar estas categorias,
- c) O conhecimento que é expresso e transmitido por estas unidades (Cabré, 2004b: 90).

Quando nos debruçamos sobre a localização como processo, vemo-nos, assim, perante a necessidade de ter presente na sua análise e no seu estudo os contributos da terminologia e da sua abordagem teórica. Se na análise do processo de localização nos deparamos com a questão do modo como o conhecimento especializado é transmitido, a resposta surge, a nosso ver, colocando a terminologia como o modo privilegiado para essa transmissão, graças à sua profundidade de análise da língua de especialidade, sobretudo em situações de comunicação entre especialistas.

A terminologia constitui-se, assim, como um elemento crucial sempre que a informação e o conhecimento relativo a um domínio de especialidade são gerados, utilizados, armazenados e processados, transmitidos, implementados, traduzidos, localizados ou interpretados. De facto, uma terminologia, enquanto conjunto de termos específicos a um domínio, estruturados, inter-relacionados e que se constituem como elementos de organização e designação do conhecimento, funciona como pré-requisito multifuncional para e como factor de comunicação intercultural (Budin, 2005). Por outro lado, o recurso às metodologias terminológicas permite capturar de forma mais profunda a

⁵⁰ A tradução constitui um fim em si mesma. A terminologia, enquanto conjunto de unidades especializadas de um domínio de comunicação técnico-profissional não constitui *per se* um produto final de comunicação – constitui antes um meio para levar a cabo outras actividades de carácter linguístico (Cabré, 2004b: 91).

relatividade intercultural e muitas das assimetrias existentes entre as representações de conhecimento específicas a cada cultura.

Decorre daqui a sua importância e centralidade para o desenvolvimento do processo de localização e para a garantia da qualidade e consistência da localização, evidenciada por diferentes autores, como por exemplo Schmitz (2007), para quem a terminologia constitui o elemento primário de transferência de conhecimento entre o produtor e o utilizador final, ou por Quirion (2003: 548), que vê a terminologia como o elemento nuclear da localização, dado que a terminologia relacionada com o produto deve, de acordo com a sua opinião, ser estabelecida tendo em conta os *“linguistic usages of the local specialists (horizontal axis) and ensuring rigorous terminological consistency between all the versions [of the] product and with related articles (vertical axis)”*.

Gouadec (2003: 528), por seu lado, associa o processo de gestão da terminologia ao da internacionalização, e afirma que, idealmente, os produtos deveriam ser internacionalizados desde o momento da sua concepção de modo a facilitar o trabalho terminológico. Esta perspectiva cria pontos de contacto entre o processo de localização e o processo terminológico, dado que uma terminologia procura, através de uma análise sistemática, uma estruturação e representação de um domínio do saber, e a localização deve, à partida, procurar considerar e gerir todos os elementos respeitantes ao produto que se enquadrem no âmbito das suas especificidades culturais, linguísticas, técnicas, religiosas, filosóficas, do sistema de valores e, finalmente, dos métodos de apresentação. (Gouadec, 2003: 528).

Esta proximidade é reforçada pela existência de um conjunto de recursos, como bases de dados terminológicas e dicionários técnicos mono e multilingues, entre outros, que apoiam o localizador no acesso à informação especializada e no seu processo de consulta e decisão. Finalmente, a compilação de dados a partir de *corpora*, sobretudo os extraídos a partir de *corpora* paralelos ou comparáveis, pode resultar na construção de recursos que contribuam para um acesso à informação mais rico e, possivelmente, de carácter mais actual. Da conjugação destes elementos resulta uma relação simbiótica entre o estudo da terminologia e o estudo da localização que coloca a terminologia no centro do processo de localização, terminologia que Warburton (2010) considera como *o ADN da globalização e da localização*.

De facto, para o bom desenvolvimento, disseminação e uso de um produto, a terminologia tem que ser estabelecida tendo em conta os usos linguísticos dos especialistas (comunidades) locais e a necessidade de assegurar consistência terminológica entre todas as versões (de um conteúdo digital). A boa gestão terminológica pode, assim, ser vista como um contributo essencial para a boa gestão do processo de localização, constituindo-se o trabalho terminológico como um elemento nuclear a todo o processo de localização. Como explicita Karsch (2009: 123), a terminologia normalizada é o suporte principal dos processos de tradução e localização orientados para a qualidade.

Nas palavras de Karsch (2009: 144), cenários mais ou menos complexos de localização estão profundamente dependentes de uma boa gestão da terminologia. Para a autora, de modo a maximizar o retorno no investimento, a gestão sistemática da terminologia deve ser baseada numa análise sistemática dos conceitos, podendo resultar desta, eventualmente, ontologias formalizadas. Adverte, no entanto, para a dificuldade em satisfazer os pré-requisitos que garantam o sucesso quer da construção das terminologias quer de sistemas baseados em ontologias, uma vez que é necessário interligar processos, ferramentas e pessoas num ambiente altamente complexo de apoio à construção daqueles artefactos.

A localização não é, assim, apenas uma questão de língua. É um processo que envolve um conjunto de actividades e actores com a preocupação de possibilitar a comunicação entre diferentes línguas a partir da correcta gestão e transferência de informação e conhecimento efectuadas com base na análise sistemática dos conceitos de um domínio de especialidade, nomeadamente enquanto processo integrante da gestão do conhecimento baseado em ontologias.

Enquanto processo integrado na construção de uma ontologia, a localização surge ligada a dois elementos essenciais a esta construção: a terminologia e a conceptualização. No que respeita à conceptualização, algumas ontologias, quando localizadas, podem ser alvo de um processo de reconceptualização e necessitar de adaptar a sua estrutura conceptual de modo a melhor enquadrar os conceitos⁵¹ de uma comunidade linguística e cultural específica. No que concerne à terminologia, esta desempenha um papel decisivo no

⁵¹ “Thoughts of reference” no triângulo de Ogden and Richards.

processo de localização, uma vez que está relacionada com as designações atribuídas aos diferentes elementos da ontologia e é o veículo que permite que os termos da ontologia possam ser mais correctamente expressos em mais do que uma língua.

A terminologia e a conceptualização estão interligadas, uma vez que qualquer modificação⁵² produzida na representação conceptual poderá ter um reflexo sobre a terminologia em uso e conduzir a alterações, alterações essas que, na nossa perspectiva, podem ser motivadas pelo desenvolvimento do processo de localização.

É nosso entender que, no contexto de trabalho bilingue e multilingue em que se insere a localização de ontologias, e tendo em conta a disponibilidade e o volume sempre crescente quer de *corpora* quer de outros recursos linguísticos e conceptuais, muitos dos quais multilingues, a terminologia pode desempenhar um papel importante ao assistir, com base nas metodologias e nos recursos que disponibiliza, os terminólogos, tradutores e especialistas nas tarefas de acesso ao conhecimento, na descrição conceptual, na criação do sistema de conceitos, na formulação de definições e, finalmente, no estabelecimento de equivalentes entre os termos entre as diferentes línguas.

Importa, assim, associar o trabalho terminológico ao de localização, devendo, na nossa perspectiva, o trabalho terminológico ter como ponto de partida uma perspectiva conceptual, de base onomasiológica, que parta dos conceitos e que permita que a localização não se concentre, apenas, na tradução dos nomes, mas antes na procura das denominações que, em cada língua, melhor descrevem um conceito especializado.

2.3. Localização de ontologias

A engenharia de ontologias vê a localização de ontologias como um subtipo da localização de *software*, na qual o produto é uma especificação de uma conceptualização partilhada de um domínio particular, i.e., uma ontologia, produto que será utilizado por uma

⁵² Para Montiel-Ponsoda (2011: 198), esta modificação da representação conceptual acontece sobretudo “*in culturally-influenced domains, whereas internationalized domains only require modifications in the terminological layer with no impact on the conceptualization, since the latter is valid and shared among the cultures involved*”.

aplicação específica. Neste contexto a localização de ontologias foi definida como sendo a adaptação de uma ontologia a uma língua e cultura particular (M. C. Suárez-Figueroa & Gómez-Pérez, 2008).

Esta definição foi revista em 2010 por Cimiano *et al.*, com o intuito de propor uma definição mais abrangente com o objectivo expresso de retirar a ênfase colocada na necessidade *de adaptação a uma língua alvo* (2010: 01), a seu ver, limitadora.

Cimiano *et al.* (2010: 01) definem a localização de ontologias como sendo o processo de adaptação de uma dada ontologia às necessidades de uma comunidade particular, que pode ser caracterizada por uma língua comum, uma cultura comum ou um ambiente geopolítico específico⁵³.

Segundo os autores, a adaptação à língua da comunidade alvo representa apenas um dos aspectos possíveis da localização de ontologias, o que leva a que os autores descrevam este processo como um tipo específico de actividade de reengenharia de ontologias onde os requisitos (de reengenharia) resultam das necessidades da comunidade para a qual a ontologia será adaptada.

Tal como descreve Montiel-Ponsoda (2011: 189), a primeira definição é mais genérica e refere-se à *língua e cultura* como uma entidade única *“in the sense that any differences in the categorization of reality because of cultural discrepancies will be inevitably mirrored in the language”*. A segunda definição coloca a ênfase na adaptação da ontologia às necessidades de uma comunidade alvo que pode ser caracterizada:

(a) by speaking a different language from that of the original community of users, or (b) by speaking the same language of the original community of users, but belonging to a different culture or geopolitical environment.

Em qualquer dos casos, as definições percebem a localização de ontologias como um processo de adaptação de uma ontologia e da sua documentação em língua natural às necessidades dos utilizadores alvo.

Para Montiel-Ponsoda (2011: 190), quer a abordagem à localização de *software* quer a abordagem à localização de ontologias, seguem uma orientação muito pragmática para o

⁵³ No original: *“Ontology localization is the process of adapting a given ontology to the needs of a certain community, which can be characterized by a common language, a common culture or a certain geopolitical environment”*.

desenvolvimento do processo, uma vez que procuram reutilizar aplicações de *software*, ontologias já existentes e outros recursos semânticos multilingues para desenvolver o processo de localização. Em ambas as abordagens, o ponto de partida é um “produto” criado numa determinada cultura e numa determinada língua, i.e., um produto monolíngue.

No entanto, na nossa opinião, esta visão nem sempre corresponde à realidade, uma vez que a criação de uma ontologia pode não seguir um paradigma de reutilização de recursos, por um lado, e a construção da representação conceptual do conhecimento do domínio pode decorrer sem ter em conta outras ontologias, isto partindo do princípio que existem.

Por outro lado, atendendo a que cada vez mais comunidades trabalham em rede, e que os seus actores são oriundos de diferentes proveniências culturais e linguísticas, o ponto de partida para a localização – no caso de uma ontologia construída neste contexto - deixa de ser um produto criado numa determinada cultura – uma vez que estão várias em presença – passando a ser antes um produto construído a partir da intersecção entre várias culturas. Finalmente, a língua utilizada para construir a ontologia pode não ser a língua de origem dos sujeitos, mas uma língua de trabalho determinada e acordada pela comunidade, como forma de garantir a comunicação e o desenvolvimento dos processos de conceptualização e de localização.

Concordamos, no entanto, com a autora quando afirma que, na localização de *software*, o produto original é adaptado às diferentes comunidades culturais e o resultado será, normalmente, utilizado de modo independente do produto original. Já no que se refere ao contexto da construção de ontologias, a ontologia localizada pode ser utilizada de modo independente da ontologia original, ou pode também ser utilizada para apoiar uma aplicação na qual várias línguas naturais necessitem de interoperar. Neste caso, o *output* será uma ontologia multilíngue (Montiel-Ponsoda, 2011: 190).

As dimensões do trabalho e dos processos envolvidos na localização, bem como a natureza dos conteúdos a localizar, têm vindo a evoluir à medida que surgem novas ferramentas e novos desafios. Apesar de a localização ser um processo bem desenvolvido, cujas metodologias são utilizadas com sucesso crescente pela indústria da língua no desenvolvimento e adequação de conteúdos multilingues, não foi ainda suficientemente explorada como elemento de apoio à construção de ontologias multilingues. A falta de

sistematicidade na análise desta problemática relaciona-se, na nossa perspectiva, com a complexidade inerente ao artefacto a localizar – a ontologia. Tal como explicita Budin (2005: 114), o objectivo mais genérico do desenvolvimento de processos de localização de ontologias é permitir a interoperabilidade semântica interlinguística⁵⁴ em ambientes de informação de grande volume que contém, normalmente, um grande número de recursos de conhecimento heterogéneos.

A especificação de uma abordagem que potencie a interoperabilidade semântica entre estes recursos requer, segundo o autor, uma modelação detalhada das situações, “*a description framework and a methodology that allows us to focus on capturing semantic complexity, richness, and diversity in real-world resources*” (*idem, ibidem*).

Uma vez que o conhecimento e a sua representação em termos ontológicos não se restringem ao uso de uma língua natural em particular, o recurso ao multilinguismo tornou-se, assim, uma necessidade evidente na construção de ontologias. Partindo da observação à literatura existente (Budin, 2005; Cortese, 2007; Dam *et al.*, 2005; Montiel-Ponsoda, 2011) percebe-se que a preocupação crescente com os fenómenos multilingues e a sua integração no desenvolvimento de ontologias estiveram entre os primeiros problemas e dificuldades detectados na evolução da engenharia de ontologias.

A integração e o enriquecimento das ontologias com elementos multilingues como equivalentes, eventuais sinónimos ou variantes, e definições em diferentes línguas, foram, no entanto, sendo investigados de modo menos sistemático, dada a dificuldade que representam. Só mais recentemente (Aguado de Cea *et al.*, 2007; Buitelar *et al.*, 2006; Cimiano *et al.*, 2007; Fu *et al.*, 2009; Gödert, 2008; Montiel-Ponsoda, 2008, 2011) se notou uma evolução nos sistemas técnicos e abordagens teóricas disponíveis, bem como no

⁵⁴ De acordo com a *European Interoperability Framework*, descrita em <http://www.softwarepublico.gov.br/>, *interoperabilidade semântica* refere-se à capacidade que um sistema deve ter para combinar dinamicamente as suas informações com outras, recebidas de outros sistemas, e processá-las para produzir um significado. Como explicita a *European Interoperability Framework*, Version 1.0 (2004: 16) “*This aspect of interoperability is concerned with ensuring that the precise meaning of exchanged information is understandable by any other application that was not initially developed for this purpose. Semantic interoperability enables systems to combine received information with other information resources and to process it in a meaningful manner. Semantic interoperability is therefore a prerequisite for the front-end multilingual delivery of services to the user*”. A *interoperabilidade semântica interlinguística* refere-se, assim, à habilidade que um sistema deve ter para combinar dinamicamente as suas informações com outras, recebidas de outros sistemas e representadas em diferentes línguas naturais e processá-las para produzir um significado que possa ser compreendido pelos utilizadores.

desenvolvimento de alguns princípios epistemológicos que possibilitaram a abertura de novos campos de reflexão e análise nesta área.

De facto, quando acedemos a repositórios de ontologias como o DAML⁵⁵ ou o Open Ontology Repository⁵⁶, verificamos que apenas um número diminuto é bilingue ou multilingue, sendo grande parte desenvolvidas e implementadas somente com recurso à língua inglesa. Este facto pode resultar, num contexto multilingue de comunicação e acesso ao conhecimento, em barreiras linguísticas para os diversos utilizadores e em restrições no acesso ao e na utilização do conhecimento a partir de ontologias.

Daqui advém uma dificuldade acrescida para os utilizadores, as organizações e redes internacionais que lidam diariamente com informação multilingue, onde se percebe uma cada vez maior necessidade de representar o conhecimento em mais do que uma língua natural nas ontologias que operam ou venham a operar nos seus sistemas de gestão da informação, de modo a acederem ao conhecimento expresso nas diferentes línguas com maior eficácia.

2.3.1. Localização de ontologias: dimensões e dificuldades

A escassez de ontologias formalizadas que contenham, associada, informação linguística em mais do que uma língua natural provém de vários factores, a começar pela dificuldade inerente à escolha das metodologias a adoptar para a conceptualização e representação do conhecimento num ambiente de desenvolvimento e localização de ontologias multilingues e pelo número diminuto de ferramentas específicas que apoiem este processo.

Uma parte das dificuldades identificáveis em qualquer sistema de localização reside em resolver problemas que designaremos de tradicionais e que decorrem do processo de tradução, como a inexistência de equivalentes para designar alguns conceitos na língua de chegada, a existência de termos polissémicos e de quasi-sinónimos, ou problemas de variação linguística, entre outros.

⁵⁵ <http://www.ai.sri.com/dam/ontologies/>

⁵⁶ <http://ontolog.cim3.net/cgi-bin/wiki.pl?OpenOntologyRepository>

Como reconhece Rogers (2007b: 15), nem sempre é fácil atingir, dentro de um domínio específico, uma relação de um para um (*Eineindeutigkeit*) entre termo-conceito e conceito-termo, noção que, segundo a autora já estaria presente na percepção de Wüster (1985: 79), para quem esta possibilidade, aplicada em larga escala, representaria *ein frommer Wunsch*.

Outras dificuldades, mais específicas, advêm sobretudo dos problemas de ordem linguística que surgem na associação dos significados de termos de diferentes línguas a conceitos representados numa ontologia, uma vez que não se pode dizer que os significados dos conceitos são coincidentes, dado que os primeiros estão, não raramente, relacionados com uma visão particular de uma língua e de uma cultura, ao passo que os conceitos de uma ontologia procuram capturar objectos do mundo real, sendo definidos e organizados formalmente de acordo com critérios acordados por consenso entre especialistas (Montiel-Ponsoda *et al.*, 2008: 69). Tal como afirma Roche (2007), só podemos comunicar e partilhar informação “*que si l’on s’accorde un tant soit peu sur les mots utilisés et leur signification*”.⁵⁷

Para além das dificuldades descritas, o processo de localização de uma ontologia, artefacto específico para a representação do conhecimento de um domínio e construído num contexto e com um propósito particular, tem que atender a outros requisitos, dos quais se destacam os relacionados com:

1. a definição e delimitação do domínio ou subdomínio(s) a conceptualizar;
2. a tipologia, classificação e conversão de recursos semânticos a adaptar para tornar o processo célere e enriquecer a ontologia;
3. a simultaneidade entre a conceptualização e o processo de localização
4. o desenvolvimento de metodologias e ferramentas necessárias à integração entre os processos de conceptualização e localização;
5. a limitação temporal normalmente imposta aos processos de conceptualização e de localização;
6. o modo de integração e (re)utilização de recursos e ferramentas linguísticas já disponíveis.

⁵⁷ Para o autor, o significado dos termos repousa sobre uma representação partilhada e consensual de uma modelação de um domínio de aplicação, constituindo esta representação uma ontologia.

Para Espinoza *et al.*, (2009: 34) é necessário ter em conta outras dimensões no processo de localização de ontologias, nomeadamente, as dimensões:

1. da tradução,
2. da gestão dos elementos multilingues ao longo do ciclo de vida da ontologia,
3. do modo de representação formal do multilinguismo na ontologia.

Assim, no que respeita ao que designam como *processo de tradução*, e uma vez que cada cultura percepçiona o mundo de modo distinto, ao localizar os elementos multilingues de uma ontologia deparamo-nos, segundo os autores, com diferentes situações. A saber:

1. *existência de um equivalente exacto*: esta equivalência exacta é sobretudo mais frequente em domínios técnicos altamente especializados, onde se pode encontrar uma equivalência directa entre termos de línguas diferentes que designam certos objectos ou processos.
2. *existência de equivalentes dependentes do contexto*⁵⁸: quando um termo numa língua pode ser traduzido por mais do que um equivalente na língua alvo, cabendo ao utilizador escolher o mais adequado tendo em conta o contexto da ontologia, o registo sociolinguístico em que a ontologia será usada, etc.
3. *existência de incompatibilidade entre as conceptualizações*: quando uma cultura categoriza a realidade com um grau de granularidade que não encontra correspondência na língua alvo, ou quando esta realidade não existe na língua alvo. Esta situação pode resultar numa lacuna lexical ou na falta de equivalentes na língua alvo.

Uma outra dimensão identificada é relativa aos problemas de gestão da ontologia. De facto, para além das dificuldades inerentes à localização dos termos da ontologia *per se*, a manutenção e actualização da ontologia localizada, ao longo do seu ciclo de vida, requer, de igual modo, uma atenção especial. As dificuldades colocadas por este processo não foram ainda totalmente consideradas e desconhecemos a existência de metodologias específicas para a gestão das possíveis alterações dos conceitos representados nas ontologias. Da

⁵⁸ Segundo os autores o contexto está directamente ligado ao contexto de construção e de uso da ontologia e ao processo de escolha do equivalente mais adequado, dado que, na nossa perspectiva, não podemos ligar os equivalentes dependentes de contexto à existência de diversidade na denominação.

análise feita aos trabalhos sobre gestão de ontologias, não encontramos nenhuma proposta que abordasse directamente a problemática da gestão da informação multilingue.

A gestão dos elementos multilingues de uma ontologia coloca-nos perante diferentes cenários, nomeadamente a:

- i. adição de um termo à ontologia, que obriga a que seja traduzido para todas as línguas representadas na ontologia,*
- ii. remoção de um termo, que obriga à remoção de todas as suas traduções,*
- iii. atribuição de uma nova designação a um conceito, obrigando à revisão de todos os termos relacionados nas restantes línguas (Espinoza et al., 2009: 34).*

Estas alterações implicam uma atenção específica a este processo, bem como a disponibilidade de recursos e ferramentas linguísticas, por um lado, e de especialistas, por outro, para acompanhar, validar e apoiar a sua execução.

A metodologia de representação das diferentes línguas naturais na ontologia coloca, também ela, dificuldades. O processo de localização dá origem a uma ontologia com termos em diferentes línguas, tendo Montiel-Ponsoda (2011: 204) identificado três processos, que vê como os mais utilizados para o desenvolvimento da localização, e que descrevemos em seguida:

Processo 1: Including multilingual data in the ontology meta-model: this implies localization at the terminological layer⁵⁹ since the ontology conceptualization remains unmodified.

Este processo, cujo uso é recorrente na comunidade da engenharia de ontologias, é baseado nas duas linguagens mais utilizadas para o desenvolvimento de ontologias: *rdf* e *owl*. A inclusão dos termos na ontologia é feita através das propriedades *rdfs:label* e *rdfs:comment* e permite associar formas de palavras e descrições aos elementos da ontologia.

⁵⁹ Montiel-Ponsoda (2011: 198) distingue entre camada terminológica “*terminological layer*” e camada conceptual “*conceptual layer*”. Esta distinção é feita com base em Barrasa, 2007, e procura identificar, no seio de uma aplicação de construção de ontologias, as diferentes “camadas” relacionadas com o processo de explicitação formal de uma ontologia. Para Barrasa (2007), uma ontologia, em termos da sua implementação computacional, divide-se em seis camadas:

- I. Lexical layer: characters and symbols that make up the syntax (ASCII encoding, UNICODE, etc.);*
- II. Syntactic layer: structure of characters and symbols, i.e., the grammar. It embraces different representation languages (e.g. RDF(S), OWL, etc.);*
- III. Representation paradigm layer: paradigm followed in the representation of the ontology (frames, semantic networks, Description Logics, etc.) that allows for certain ways of expressing and structuring knowledge;*
- IV. Terminological layer: terms or labels selected to name ontology elements;*
- V. Conceptual layer: related to conceptualization decisions, such as granularity, expressiveness, perspective, etc.;*
- VI. Pragmatic layer: final layout of the model according to the user’s needs.*

O recurso a este processo permite que os termos sejam incluídos na ontologia em tantas línguas quantas as desejadas pelo utilizador. Este processo oferece, na perspectiva da autora, a possibilidade de uma representação adequada para ontologias de domínio altamente especializadas, sobretudo em domínios *internacionalizados* ou *normalizados*, uma vez que o tipo de conhecimento partilhado entre as diferentes comunidades linguísticas e culturais e a relação de equivalência entre os termos das diferentes línguas são adequados.

No entanto, a sua utilização impõe restrições à quantidade de informação linguística a disponibilizar, estando esta limitada a sequências (*strings*) que não contêm qualquer informação quanto ao significado do termo na língua de chegada, ou quanto à origem da informação, o que pode dificultar o processo de localização para as diferentes línguas naturais, assumindo-se, por outro lado, a equivalência exacta entre os termos nas diferentes línguas, o que, como descrevemos já, nem sempre acontece na realidade.

Processo 2: Combining the ontology meta-model with a mapping model: this allows localization at the conceptual layer since conceptualizations in different languages are mapped to each other. Linguistic information is also included in the ontology. The mappings establish links or equivalence relations among the various conceptualizations.

Este processo de modelação do conhecimento é organizado a partir das estruturas conceptuais específicas a uma dada comunidade de utilizadores, sendo possível o uso de mais do que uma opção à integração da informação multilingue nas ontologias, nomeadamente através de:

1. *Binary mappings in an orthogonal graph. In this case, each monolingual ontology organizes knowledge of a certain culture, and is mapped to the rest of ontologies in a pair-wise fashion.*
2. *Binary mappings in a radial graph. In this option, monolingual ontologies are mapped to each other through an interlingua consisting of a set of common concepts that allow establishing equivalences. (Montiel-Ponsoda, 2011: 206)*

Estas opções permitem, na perspectiva da autora, a criação de uma conceptualização independente em cada língua e cultura envolvidas no processo e o estabelecimento de mapeamentos entre as diferentes conceptualizações, em que cada conceptualização diz respeito à categorização da realidade tal como é percebida por uma dada cultura, o que

pode enriquecer a ontologia em determinados domínios. Obriga, no entanto, a um grande esforço, quer em termos do desenvolvimento das diferentes conceptualizações quer da sua interligação/mapeamento, não disponibilizando, por outro lado, uma quantidade de informação linguística muito superior à abordagem anterior.⁶⁰

Processo 3: *Associating the ontology with an external linguistic model: localization is performed at the terminological layer, which is now represented by a complex external model that stores linguistic information. (Idem, ibidem)*

Nesta proposta, a associação de informação multilingue à ontologia acontece através da ligação a dados linguísticos armazenados fora da ontologia, por exemplo, numa base de dados⁶¹. Para a autora, a vantagem deste tipo de representação reside no facto de permitir o enriquecimento das ontologias de domínio com o que descreve como sendo *modelos linguísticos* ricos e complexos. Estes correspondem, no fundo, a aplicações externas à ontologia, desenvolvidas e utilizadas para processar e armazenar a informação linguística de forma independente. Na perspectiva da autora, uma vez que estas aplicações correspondem ao que denomina de modelos externos portáteis (*external portable models*), estes podem ser associados a qualquer ontologia de domínio.

Esta proposta levou ao desenvolvimento de diferentes ferramentas linguísticas, como o *Linguistic Information Repository* (LIR) ou o *LingInfo*, que descreveremos adiante. A principal vantagem deste processo é a de não requerer a criação de novas conceptualizações do domínio e a de permitir a reutilização de ontologias já existentes na *Web*, ainda que tal possa conduzir à perda de especificidade ou à não representação das especificidades de algumas línguas, uma vez que estamos perante uma única conceptualização.

A decisão quanto à utilização de cada uma dos três processos acima descritos depende de dois factores: 1) o tipo de conhecimento de domínio representado na ontologia e 2) a quantidade de informação linguística necessária ao funcionamento da aplicação final.

⁶⁰ Apesar das dificuldades associadas ao recurso a esta abordagem, há um interesse cada vez maior no seu uso, tal como descrito em Euzenat *et al.*, 2009, interesse esse que advém do número crescente de ontologias construídas em diferentes línguas que descrevem o mesmo domínio de conhecimento e que seria interessante tornar interoperáveis.

⁶¹ O modelo para representar e organizar a informação linguística pode ser uma base de dados (como no caso do GENOMA-KB ou OncoTerm), ou uma ontologia (como no caso do LingInfo, LexOnto ou AGROVOC Concept Server). (Montiel-Ponsoda *et al.*, 2008).

Uma ontologia é uma *especificação formal de uma conceptualização partilhada*, obtida através de negociação e consenso, que representa e organiza explicitamente o conhecimento, mas, como alertam Espinoza *et al.* (2009), certas conceptualizações possuem uma tendência para reflectir particularidades culturais. O domínio a categorizar torna-se, assim, um elemento crítico para o desenrolar do processo de localização.

Tal como afirmam os autores, se a conceptualização é partilhada por todas as culturas implicadas na actividade de localização, a localização afectará apenas o que designam por camada terminológica. No entanto, se o enfoque recair sobre domínios culturalmente dependentes, como o jurídico, no qual as categorizações tendem a reflectir particularidades de uma cultura específica, a localização pode afectar directamente a conceptualização.

Tendo em conta, por um lado, esta perspectiva e, por outro, o domínio e a quantidade de informação linguística necessária à localização, Espinoza *et al.* (2009) descrevem dois processos para a integração de informação multilingue nas ontologias:

1. Se a representação conceptual do domínio for consensual, apresentam duas opções: optar pela inclusão da informação multilingue na ontologia, ou pela associação de um modelo externo de base linguística (ex.: *LingInfo*) ao processo de localização⁶².
2. Se a representação conceptual do domínio reflectir particularidades de uma cultura específica e se se percepcionar como necessária a representação das diferenças culturais, os autores apresentam, de novo, duas opções: a criação de uma conceptualização por língua e cultura envolvidas, ou a associação de um modelo externo (ex.: *LIR*) que permita dar conta das divergências culturais ao nível da camada terminológica.

Os dois processos propostos são, na nossa perspectiva, demasiado estanques, uma vez que as fronteiras entre uma representação consensual de um domínio e uma representação dependente das culturas nem sempre são nítidas, sobretudo no caso de uma construção partilhada de uma conceptualização que envolva diferentes comunidades

⁶² A decisão entre estas duas opções dependerá das necessidades linguísticas da aplicação a desenvolver. “*If morphosyntactic data is needed for the purpose of Information Retrieval or Information Extraction, for example, the most suitable option will be the association of an external model such as LingInfo, which enriches the ontology with a great amount of morphosyntactic information.*” (Espinoza *et al.*, 2009)

linguísticas e culturais. Por outro lado, não prevê a existência de domínios onde uma parte da conceptualização possa ser consensual e outra marcadamente cultural, de modo a representar elementos específicos a determinadas culturas.

2.4. Abordagens ao processo de localização de ontologias

A necessidade de construir os elementos conceptuais que constituem as ontologias com base em línguas naturais tornou-se, segundo Falk *et al.* (2009: 02), evidente, tendo em conta que as ontologias são desenvolvidas por indivíduos diferentes, provenientes de comunidades distintas, durante um período longo de tempo e são pensadas para serem reutilizadas, o que implica o recurso à informação linguística e terminológica de modo a estabelecer e assegurar a consistência dos sistemas linguístico e cognitivo humano com o sistema computacional de base ontológica.

A utilização de ontologias representadas em várias línguas torna-se, assim e cada vez mais, uma componente integral de uma estratégia de gestão eficiente de ontologias, estando esta componente ligada aos esforços de normalização, permitindo uma economia de esforço na sua reutilização e na sua manutenção.

Os problemas e dificuldades supramencionados, que retratam e oferecem uma visão da especificidade e complexidade do artefacto e dos conteúdos a localizar, bem como da riqueza de possibilidades abertas ao desenvolvimento de aplicações e abordagens, conduziram ao desenvolvimento de projectos e ao surgimento de propostas de características distintas que apresentam modelos, metodologias e ferramentas com o intuito de contribuir para o estabelecimento da interligação da informação terminológica e linguística multilingue com a ontológica.

As metodologias e ferramentas que analisámos a seguir foram desenvolvidas com objectivos distintos e partem de pressupostos teóricos muito diversos. No entanto consideramos importante a sua identificação e análise, dados os contributos relevantes, que resultaram da sua implementação e desenvolvimento, para a integração da informação terminológica e linguística multilingue nas ontologias.

O nosso enfoque recairá nas metodologias propostas pelos projectos Eurowordnet (Vossen, 1997, 1998)⁶³, GENOMA-KB (Cabré *et al.*, 2004) e Termontography (Kerremans & Temmerman, 2004), nos modelos LinInfo (Buitelaar *et al.*, 2006) e LIR - Linguistic Information Repository (Peters *et al.*, 2008) e na ferramenta LabelTranslator (Espinoza *et al.*, 2008)⁶⁴. Este enfoque teve em conta a relevância e o contributo obtido a partir do seu desenvolvimento para o processo de especificação multilingue de ontologias.

Uma das abordagens que mais destaque alcançou neste domínio é a proposta no projecto Eurowordnet (Vossen, 1998). O Eurowordnet é uma base de dados lexical multilingue, que não é específica a qualquer domínio, construída com base nos Wordnets existentes, (representando cada um deles uma língua europeia diferente), que propõe um modelo constituído pelo que os autores denominam de ontologias monolingues autónomas (Peters *et al.*, 1998:150), cada uma das quais reflectindo as especificidades linguísticas e culturais de uma determinada língua, ligadas entre si por um *Inter-Lingual-Index*, que permite o acesso às traduções e o estabelecimento de um mapeamento através de equivalências entre as diferentes ontologias.

Tal como no Wordnet, a informação sobre os nomes, verbos, adjectivos e advérbios é ligada por meio de *synsets*, definidos como um conjunto de “*words with the same part-of-speech that can be interchanged in a certain context*” (Vossen, 2004). Os *synsets* são relacionados entre si por relações semânticas como a hiponímia e a meronímia, entre outras, tal como representado na figura que se segue.

⁶³ Apesar de não ser considerada uma ontologia por diversos autores (Wright e Budin, 2001), posição com a qual concordamos, tomámos a decisão de analisar as abordagens propostas no âmbito do projecto EuroWordnet, dada a relevância dos resultados obtidos e, posteriormente, a sua grande reutilização em projectos de construção, alinhamento e localização de ontologias.

⁶⁴ Foram, ao longo da análise do estado da arte, identificados outros modelos, projectos e abordagens dedicados ao enriquecimento multilingue de outro tipo de recursos, como as propostas de localização de glossários (ex.: FAOTERM), de localização de bases de dados terminológicas (ex.: Eurodicautom), de localização de thesaurus (ex.: Agrovoc; Eurovoc) ou a aplicação OntoLing (Pazienza *et al.*, 2006) que, apesar da sua importância, não serão, nesta dissertação, alvo de análise, uma vez que não se relacionam directamente com o processo de localização de ontologias.

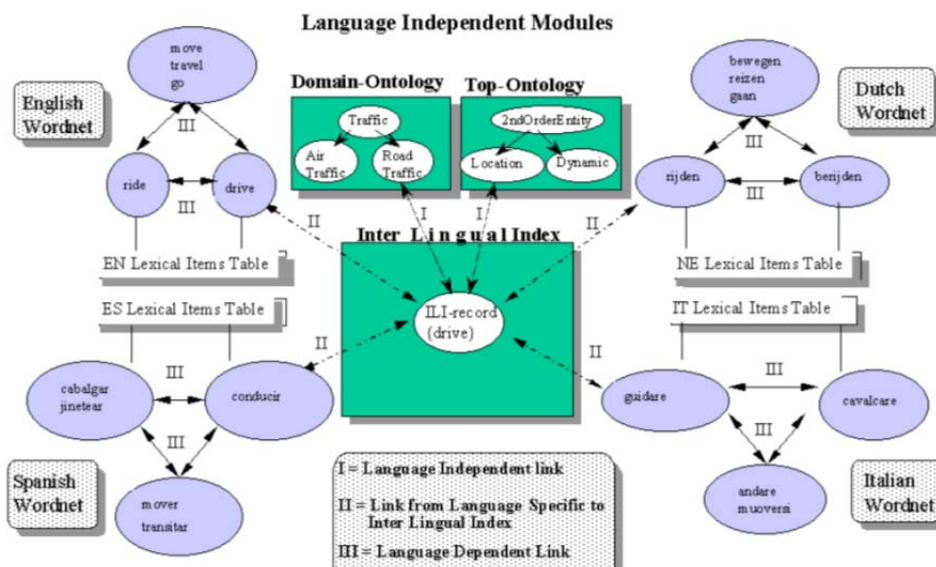


Fig. 3 – Arquitectura global da base de dados do EuroWordNet (Vossen, 2004)

Algumas das limitações inerentes a esta proposta são identificadas por Wright e Budin (2001: 889), quando afirmam que, tal como no caso de todos os *“interlingua-related resources, any automatic agent that would use this kind of hybrid resource must be committed to the constraints established by the mapping procedures”*. De facto, o uso desta abordagem implica um grande esforço, sobretudo em termos do desenvolvimento dos sistemas de mapeamento entre ontologias.

Os autores vão mais longe e identificam esta ressalva como reflectindo a experiência, comum a terminólogos e a criadores de tesouros, de que os sistemas de conceitos numa língua (L_1) raramente se comportam como um espelho perfeito de uma conceptualização similar numa outra língua (L_2). Para Wright e Budin (2001: 889)⁶⁵, esta percepção leva-os a afirmar que não se pode assumir que ontologias que se baseiam apenas no simples “senso comum” possam ser aceites como universais para outras línguas e culturas, dado que este “senso” pode ser dependente da língua ou da cultura.

Kerremans e Temmerman (2003), por seu lado, partem de uma abordagem de base terminológica e propõem uma metodologia que descrevem como funcional, a *Termontography*, que sumarizam do seguinte modo:

First of all, in close collaboration with specialists of the domain of interest, a framework of domain specific categories and intercategory relationships is developed. This framework

⁶⁵ Os autores não reconhecem o EuroWordNet como uma ontologia.

facilitates the manual and semi-automatic extraction of knowledge from a textual corpus. It will gradually evolve towards an enriched and more fine-grained network of semantic relations, reflecting culture-specific conceptualizations. This analysis results in a termontological database. (Kerremans et al., 2003: 4)

Propõem, assim, uma abordagem multidisciplinar, em que combinam as teorias e métodos da análise terminológica multilingue, partindo da perspectiva sociocognitiva (Temmerman, 2000), com os métodos e princípios para a análise ontológica propostos por Gómez-Pérez *et al.* (1996); Fernández *et al.* (1997) e Sure and Studer (2003), cujo ponto de partida é uma “pre-defined, language-independent framework of domain-specific UoUs⁶⁶ and intercategory relationships, set up in collaboration with field specialists, to which lexicalisations (from a domain-specific corpus) are mapped.”

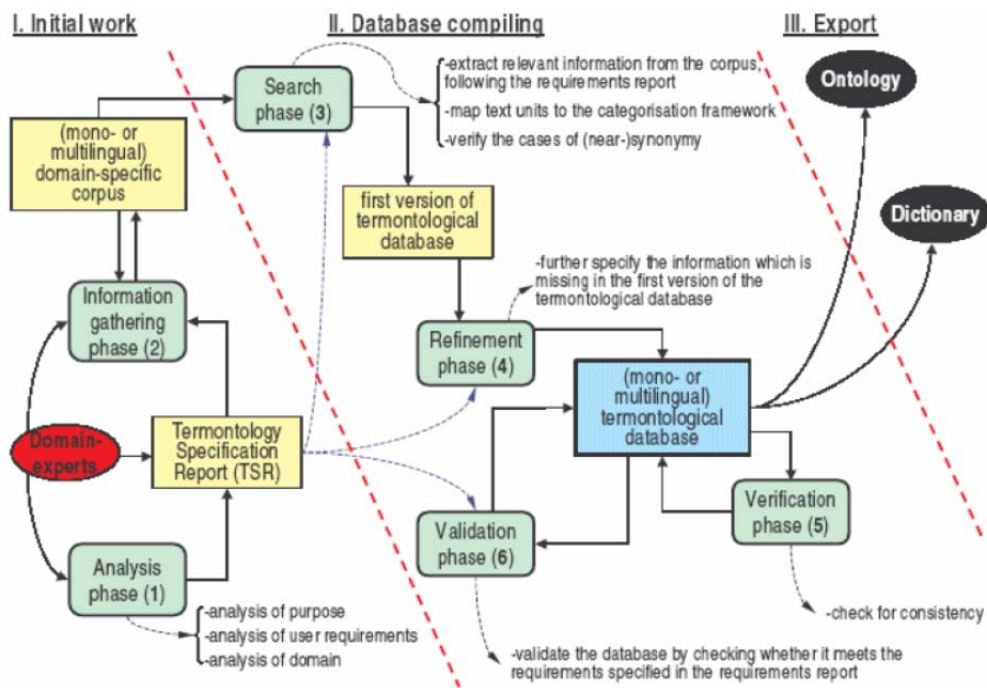


Fig. 4 – Fluxo de trabalho do modelo Termontography (2004a: 05)

Definem esta abordagem como sendo terminológica, sendo a terminologia multilingue, tal como é representado na figura anterior, extraída a partir de um *corpus* textual e estruturada tendo em conta uma “culture independent and task-oriented framework of domain-specific knowledge which can be further refined in a culture-specific

⁶⁶ UoU – Unit of Understanding – a noção de unidade de compreensão foi introduzida pela abordagem sociocognitiva à Terminologia, proposta por Temmerman (2000).

layer” (*idem, ibidem*). Esta metodologia resulta da colaboração entre terminólogos e especialistas, colaboração que os autores consideram como essencial, estando os especialistas envolvidos sobretudo numa fase inicial do processo de representação do conhecimento.

Uma outra metodologia, de base terminológica, foi proposta por Cabré *et al.* (2004) a partir do desenvolvimento do projecto GENOMA-KB. Este projecto foi desenvolvido para recolher e representar o conhecimento no domínio da genética e pode ser descrito, sumariamente, como uma base de dados de conhecimento multi-módulo, composta por quatro módulos interdependentes – o módulo da ontologia, o módulo da base de dados terminológica, o *corpus* (multilingue) e o módulo das entidades, tal como é representado na figura seguinte.

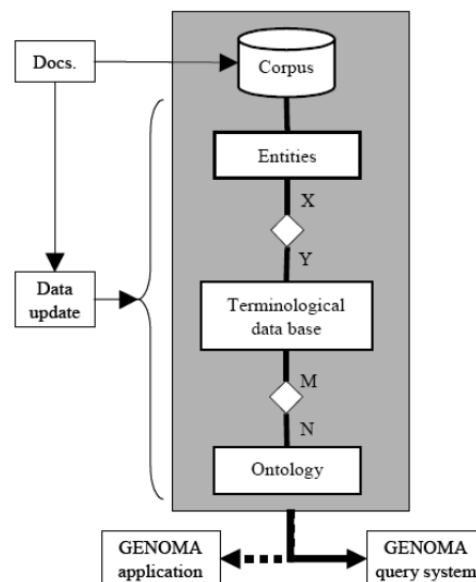


Fig. 5 - Estrutura geral da metodologia GENOMA-KB

O modelo propõe uma abordagem de cariz marcadamente terminológico, em que as *unidades de conhecimento especializado* (Cabré *et al.*, 2004) extraídas dos documentos que compõem o *corpus* multilingue são depositadas numa base de dados terminológica e ligadas a conceitos na ontologia. No módulo da ontologia, que aparece directamente ligado ao da base de dados, estes conceitos são, posteriormente, alvo de uma organização do conhecimento baseada num conjunto de relações conceptuais do tipo hierárquico e não-hierárquico.

Este modelo foi desenvolvido tendo em conta o ponto de vista defendido por Cabré *et al.* (2004: 02), segundo o qual o trabalho terminológico lida com:

specialized knowledge nodes represented by the lexical units in a specialized text (Cabré, 2002). These terminological units have clear-cut borders in each specialized domain and they correspond to a concept placed in an ontology. From this approach, we adopted a tool for managing together the terminological units used in the genetics domain and their correspondent concepts in the ontology.

Esta metodologia obriga à existência prévia da ontologia, construída recorrendo a um conjunto de base de conceitos de mais alto nível a partir da perspectiva de especialistas, para que a ligação possa acontecer, o que implica, por exemplo, em termos teóricos, a biunivocidade entre termos e conceitos e, em termos práticos, que a ontologia seja construída num momento anterior ao da base de dados⁶⁷.

Esta abordagem, que separa os diferentes componentes (base de dados, *corpus*, ontologia), tem por objectivo permitir a integração de toda a informação terminológica necessária, bem como o estabelecimento de relações entre os elementos terminológicos, o que resultaria num enriquecimento dos elementos conceptuais, sendo, a nosso ver, a abordagem que mais assumidamente segue uma via terminológica no seu desenvolvimento., ao mesmo tempo que atribui aos especialistas um papel relevante na conceptualização inicial do domínio, perspectiva com a qual concordamos.

2.4.1. Ferramentas de base linguística⁶⁸

Outras propostas adoptam uma abordagem distinta, que consiste na associação da ontologia a um sistema ou ferramenta de gestão da informação linguística externa àquela,

⁶⁷ Para Cimiano *et al.* (2010: 06), a abordagem seguida neste projecto para a conceptualização do domínio representa um bom exemplo do que entendem por: “*an internationalized domain, reflecting the common view of all the cultures represented in the project. This is a clear example of a localization project on an internationalized domain, affecting only the lexical layer in the ontology*”.

⁶⁸ O trabalho de construção e definição das ferramentas aqui descritas, e de outras, tem em consideração a tendência assente no desenvolvimento de propostas de codificação dos recursos semânticos com o intuito de aumentar a sua interoperabilidade, que resulta do uso de normas bem disseminados e aceites. Destacam-se, das propostas existentes, normas como o TMF – Terminological Markup Language (2000), o LMF – Lexical Markup Framework (2006), ou as propostas do projecto ISLE - International Standard for Language Engineering (2000).

mantendo ambos separados. A ferramenta usada para processar e representar a informação linguística pode ser uma base de dados, como no caso anterior do projecto GENOMA-KB (Cabré *et al.*, 2004) e do OncoTerm (Faber, 2002), ou uma ontologia, como no caso do LingInfo (Buitelaar *et al.*, 2006) e LIR - *Linguistic Information Repository* (Peters *et al.*, 2008), sendo estes sistemas normalmente considerados como independentes do domínio, podendo ser interligados a qualquer ontologia de domínio.

Analisaremos em seguida as abordagens seguidas no desenvolvimento do LingInfo e LIR, que, ao contrário das abordagens da Termontography ou do GENOMA-KB, seguem apenas uma abordagem lexical à localização de ontologias, não considerando a abordagem terminológica⁶⁹. Apesar de esta perspectiva não coincidir com a que propomos no nosso método (Capítulo III), consideramos importante a sua análise, uma vez que nestas propostas a representação do conhecimento do domínio e a da informação linguística surgem claramente separadas.

Uma ferramenta que mereceu algum destaque e que esteve na base de outros como o LexOnto (Cimiano *et al.*, 2007) foi a LingInfo, desenvolvida como parte do projecto *SmartWeb2*. A perspectiva conceptual seguida no seu desenvolvimento assenta na ideia de construção de uma *camada linguística* para uma ontologia ou, mais especificamente, no que os autores descrevem como um *modelo lexical multilingue para ontologias* (Buitelaar *et al.*, 2006).

As representações linguísticas na LingInfo podem consistir num identificador de idioma, em etiquetas POS (*Part of Speech*), em dados morfológicos e dados de composição sintáctica, bem como num conjunto de dados contextuais sob a forma de regras gramaticais de N-grams. Para Buitelaar *et al.* (2006a: 03), a LingInfo⁷⁰ pode ser descrito como:

an ontology and linguistic knowledge base that provides for all other ontologies used in SmartWeb linguistic information (orthographic realizations, grammatical gender, stem and inflection) on ontology classes and properties for languages that are relevant to the SmartWeb scenario, i.e. German and English (and into some respect also French).

⁶⁹ Para além dos sistemas descritos, foram identificadas outras propostas, que estão, no entanto, ainda em desenvolvimento, pelo que não foram alvo da nossa análise.

⁷⁰ Este modelo, conjuntamente com o LexOnto e a norma LMF estiveram na base do modelo LexInfo, proposto mais recentemente por Buitelaar *et al.* (2009).

Esta ferramenta, apesar de atender aos aspectos de construção de uma ontologia multilingue, é algo complexo para ser (re)utilizado por um público não-linguista, podendo dificultar as tarefas de conceptualização e implementação da ontologia ao exigir conhecimentos linguísticos relativamente complexos.

Uma outra ferramenta, mais recente e que tem vindo a ser reutilizada no âmbito de diferentes projectos é o LIR - *Linguistic Information Repository* (Peters *et al.*, 2008). Esta ferramenta foi recentemente integrada e desenvolvida com maior profundidade no âmbito do projecto NeOn⁷¹, com o objectivo, tal como é descrito pelos autores, de localizar uma determinada categoria da ontologia de acordo com o universo linguístico e cultural de uma língua natural e capturar as especificidades da tradução entre línguas.

A abordagem seguida no seu desenvolvimento, apesar de não considerar a perspectiva e as metodologias terminológicas, é descrita como sendo holística, uma vez que, defendem os autores, a ferramenta contém e disponibiliza um conjunto de classes completas e complementares que permitem a localização de um determinado elemento da ontologia de acordo com um contexto linguístico e cultural específico, permitindo, paralelamente, que se aceda e integre a informação multilingue contida em recursos lexicais distribuídos e heterogéneos ao garantir um acesso homogéneo à informação. Este sistema é constituído por:

*Lexical Entries that in its turn consist of Lexicalizations (or word forms), Senses expressed by means of the Definition in a certain Language Context, Usage Context, Source (of any linguistic element), Notes (that can be related to any linguistic element), and relations established between Lexical Entries as, for example, "has Synonym", "has Translation" or "has Scientific Name"*⁷².

⁷¹ <http://www.neon-project.org/>

⁷² Segundo os autores, a localização realiza-se ao nível da entrada lexical e a “ontology class is connected via the *hasLexicalEntry* relation. LEs (Lexical Entries) in different languages may be linked through the *hasTranslation* relation and within the same language through synonymy or antonymy relations. Their sense is typically represented through a definition from an external resource, which may also come from a different language. In LIR it is possible to express the fact that two words are synonym in most cases by linking their sense definitions through the *isRelatedTo* relation. Finally, within LIR all lexical entries, irrespective of the language they belong to, are members of the same lexical ontology”.

O sistema LIR pretende ser um elemento independente da ontologia e procura permitir aos linguistas o desenvolvimento das suas tarefas em ambientes distribuídos (*distributed environments*) de modo a gerir e editar a informação linguística associada à ontologia. O uso deste sistema surge associada ao uso da ferramenta LabelTranslator, um *plug-in* destinado à tradução automática dos termos da ontologia, que analisaremos adiante.

A aplicação e a abordagem seguida pela LIR, a par da LexInfo (Buitelaar *et al.*, 2009), que não analisaremos aqui, dada a relativa proximidade entre ambas, são apontadas como representativos das melhores práticas a desenvolver neste domínio (Falk *et al.*, 2009), sendo, na perspectiva dos autores, passíveis de serem utilizados para ligar informação linguística e lexical aos elementos da ontologia. Em ambos os casos a representação do conhecimento do domínio e da informação linguística aparece claramente separada, continuando a existir uma articulação, que os autores descrevem como flexível, entre língua e sistema ontológico. Falk *et al.* (2009) apontam algumas diferenças entre os sistemas, argumentando, no entanto, em favor de uma integração entre as duas propostas:

LexInfo is arguably more suitable at verbalising ontologies or translating natural language queries to database or web search queries, whereas LIR would be more useful to human knowledge engineers or lexicographers for building and maintaining a domain and lexicon ontology. It is in principle possible to translate one lexicon format to the other, but the syntax – semantics mapping information of LexInfo cannot be extracted from LIR and conversely, the lexicographic aspects represented in LIR cannot be generated from LexInfo. On this note the two models are complementary.

Ambos os sistemas procedem do seguinte modo: (1) A partir da ontologia de domínio, (2) o sistema constrói o que os autores designam de uma *ontologia lexical vazia ou padrão*, os blocos principais de construção das entradas lexicais, que correspondem aproximadamente aos componentes da ontologia de domínio, ou seja, as classes e propriedades. Os elementos da ontologia de domínio estão ligados nesta fase a entradas lexicais possivelmente vazias que, por sua vez, estão associadas a elementos da ontologia correspondente. As entradas lexicais são construídas com base na análise linguística dos termos da ontologia de domínio, comentários e/ou identificadores. (3) O léxico é, então, enriquecido semi-automaticamente, através da utilização textos de relevo para o domínio e/ou de recursos lexicais externos, como acontece com o WordNet ou a Wikipedia, como

forma de procura de informação adicional (por exemplo, definições, traduções); (4) A *ontologia lexical*, tal como a designam os autores, é concluída manualmente *a posteriori*. Por outro lado, ambos os sistemas aderem explicitamente ao uso de normas já existentes, como o LMF (Francopoulo *et al.*, 2007), e implementam soluções de outros sistemas, como o LingInfo (Buitelaar *et al.*, 2006) e o LexOnto (Cimiano *et al.*, 2007).

Os aspectos que diferenciam as várias ferramentas supradescritas são determinados pelo tipo de informação terminológica e linguística multilingue que as compõem e que são disponibilizados em cada uma. Dependendo das necessidades linguísticas do utilizador final, algumas ferramentas podem considerar-se mais apropriadas do que outras. De facto, o sistema a implementar depende, em grande parte, das necessidades predefinidas dos utilizadores e das aplicações e, no caso de uma rede colaborativa, dos propósitos da rede e, em última análise, do próprio modelo adoptado para apoiar e potenciar a partilha de conhecimento e o processo de conceptualização.

Para além das ferramentas e metodologias já descritas, foram também desenvolvidas aplicações com o objectivo concreto de automatizar (até certo ponto) o processo de localização de ontologias. Dessas concentraremos a nossa atenção naquela que consideramos a mais desenvolvida, o LabelTranslator, aplicação de localização de ontologias desenvolvida por Espinoza *et al.* (2008), que pretende ser um auxiliar à localização automática de ontologias recorrendo a uma metodologia que os autores denominaram de *Ontology Localization Activity* (OLA). O sistema desenvolvido é composto por: “1) *the extraction of possible single-lemma translations from semantic and translation resources*, 2) *the disambiguation of the translation senses*, and 3) *the ranking of the translations*”. (Espinoza *et al.*, 2008: 01)

Esta aplicação consulta bases de dados lexicais, dicionários e bases de dados terminológicas⁷³ e propõe uma metodologia para medir a similaridade semântica entre o conceito representado na ontologia e o seu contexto lexical e semântico. O sistema, tal como é descrito por Espinoza *et al.* (2008: 02), é composto por seis passos, fornecendo uma tradução automática acompanhada de mecanismos de desambiguação semântica descritos como sofisticados.

⁷³ Aplica-se apenas às línguas alemã, inglesa e espanhola.

De modo a obterem uma validação interlinguística das traduções, os recursos externos do LabelTranslator foram limitados àqueles que possuem informação multilingue, como o EuroWordnet, acrescidos de outros recursos específicos do domínio, para alargar a cobertura do vocabulário. No entanto, esta ferramenta e a metodologia a ela associada têm algumas limitações como reconhecem os autores (2008: 13), nomeadamente, a elevada dependência dos serviços de tradução da tipologia de recursos disponíveis e da sua abrangência do domínio, e a impossibilidade de aprendizagem de novos padrões lexicais, o que diminui a escalabilidade da ferramenta.

De facto, na nossa opinião, o grau de abrangência e precisão dos recursos lexicais, entre outros recursos linguísticos, tem vindo a melhorar à medida que vão sendo desenvolvidos e disponibilizados. No entanto, recursos lexicais e ontologias formais são recursos muito diferentes, construídos com propósitos diferentes, sendo necessário cuidado no estabelecimento de relações entre a informação que cada um disponibiliza. Na nossa perspectiva, qualquer que seja a relação que se estabeleça, ela será sempre imperfeita, se não for mediada pela análise de um terminólogo e/ou de um especialista.

Para além destas limitações, o nome LabelTranslator é auto-explicativo. O processo de localização ocorre sempre depois de a conceptualização estar terminada e abrange apenas os termos disponíveis no sistema, permitindo *a posteriori* um acesso mais completo à informação disponível sobre o domínio, o que condiciona a capacidade dos diferentes agentes de melhorarem as estruturas de representação desse conhecimento, obrigando-os, necessariamente, a despenderem mais tempo na manutenção da ontologia.

O LabelTranslator procura, no entanto, atribuir aos utilizadores a capacidade de decisão sobre as traduções, para que a capacidade expressiva das ontologias no que respeita ao conhecimento do domínio seja a mesma nas diferentes línguas de chegada, como se pode verificar na figura seguinte.

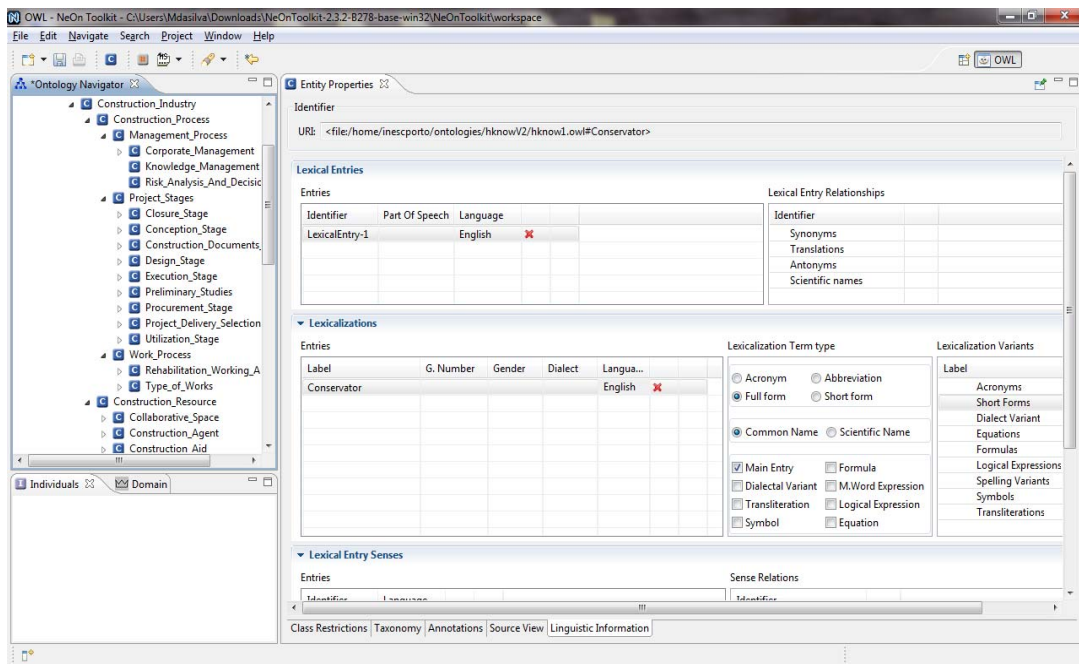


Fig. 6 – Ambiente de tradução do LabelTranslator

De modo a melhorar estas decisões, esta ferramenta foi, no âmbito do projecto Neon, associada e sincronizada com o sistema LIR, acima descrito, usando o LIR como um repositório de informação linguística.

As limitações já mencionadas e o facto de a localização, com recurso às aplicações descritas, só acontecer depois de formalizada a ontologia, levam-nos a considerar que estas ferramentas não se adequam ao fluxo de trabalho de um processo de conceptualização de uma ontologia que seja desenvolvido em mais do que uma língua no seio de uma rede colaborativa, dado que nesta participam diferentes interlocutores que precisam de recorrer a instrumentos de localização mais simples e de resultados mais imediatos que apoiem o processo de conceptualização do domínio.

De facto, o processo de localização de uma ontologia, sendo um fim em si mesmo, pode, na nossa perspectiva, constituir também um meio de auxílio à negociação e à estruturação e representação de um fluxo contínuo de informação e conhecimento do domínio, num processo que se pode descrever como bidireccional, aspectos que não são considerados nas abordagens descritas.

Por outro lado, e como resulta da nossa experiência, a qualidade dos equivalentes que advém da consulta a aplicações de tradução automática como o Google Translator é muito limitada, sobretudo em domínios de especialidade, só podendo ser considerada como

um elemento de apoio ou de consulta para o desenvolvimento do processo de localização, devendo os seus resultados ser alvo de uma revisão e validação sistemática por um terminólogo ou por um especialista de domínio.

Consideramos, ainda, que a localização dos termos que designam os conceitos representados numa ontologia vai para além da equivalência directa, como resulta do uso dos modelos de base lexical que descrevemos, sendo necessário considerar a adequação dos equivalentes obtidos ao domínio de especialidade e às representações conceptuais das diferentes línguas, tarefa que, na nossa perspectiva, não pode depender de um sistema de tradução automática ou de uma validação estatística, obrigando a um trabalho interdisciplinar que inclua não só linguistas ou tradutores, mas também os especialistas de domínio e os terminólogos.

Como reconhecem Gracia *et al.* (2012: 68) a localização dos termos de uma ontologia pode não ser suficiente para tornar uma ontologia reutilizável em diferentes cenários linguísticos e culturais. Este facto está sobretudo relacionado com o domínio representado pela ontologia, cuja conceptualização pode nem sempre corresponder às necessidades da comunidade alvo. Neste caso, tal resultaria, segundo o autor, na necessidade de modificações ou adaptações da conceptualização.

Assim, no que respeita à localização de ontologias propriamente dita, é nossa opinião que a localização de ontologias deve ser entendida, não só como o processo de tradução ou de adaptação de um termo que designa um conceito de acordo com as especificidades linguísticas e culturais do país ou região a que se destina, mas também como um processo que considera necessidade de adequação da representação dos conceitos e das suas relações semânticas ao ambiente de chegada, com respeito pela sua identidade linguística e cultural, bem como pelas suas estruturas sociais, profissionais e políticas.

Temos presente que é questionável que possamos enquadrar esta visão na sua totalidade, quando nos concentramos apenas no desenvolvimento de um sistema de conceitos correspondente a um domínio limitado do conhecimento. No entanto, consideramos que todo e qualquer domínio ou subdomínio do conhecimento é parte de um todo, para cuja construção e validação contribuem os elementos identificados.

2.5. Localização e localização de ontologias: analogias

A localização de ontologias desenrola-se num ambiente multidisciplinar, podendo envolver terminólogos, gestores de conhecimento, especialistas de domínio e engenheiros de ontologias que procuram, num contexto e com um propósito específico, estruturar e representar o conhecimento de um domínio. Sendo a localização um conceito bem estabelecido e bem aceite na comunidade de engenharia de *software*, não nos parece, no entanto, que a definição do conceito seja suficientemente abrangente para os propósitos da localização de ontologias e para a abordagem que esta dissertação pretende desenvolver.

Sobressaem, como pudemos verificar através da análise desenvolvida até este ponto, um conjunto de analogias e de diferenças a vários níveis entre o processo de localização e o processo de localização de ontologias que interessa identificar, para melhor entender como este se desenrola e, sobretudo, para melhor compreender como situar um sistema de representação do conhecimento enquanto objecto do processo de localização.

Uma primeira analogia verifica-se ao nível do objecto a localizar. De facto, um sistema de representação do conhecimento – no nosso caso, um sistema conceptual – tem, no momento da sua conceptualização, características de alguma fluidez, uma vez que é um sistema aberto à negociação e construído por consenso, sendo o seu autor, tal como num projecto típico de localização, não um indivíduo, mas um sujeito corporativo, o mesmo que inicia e, em última instância, valida a versão final.⁷⁴

Do mesmo modo, quando se localiza os resultados de uma conceptualização, expressos num sistema conceptual, os condicionalismos do localizador/tradutor provêm dos objectivos do iniciador da localização e das limitações ou orientações explícitas dadas por este, nomeadamente, em termos da delimitação do domínio e do conhecimento a representar, e não de um texto original, que não existe, ainda que possam existir outros tipos de recursos semânticos para apoiar o processo. A localização de sistemas conceptuais

⁷⁴ Esta característica é, como descreveremos mais aprofundadamente no ponto 3.7.1, importante para potenciar o processo de construção partilhada de ontologias representadas em várias línguas e de localização no seio de uma rede colaborativa.

excede assim, também ela, o processo de tradução, representando este apenas uma das fases.

Uma outra analogia, que representa, neste caso, um condicionalismo, advém do contexto e da situação comunicativa em que ocorrem a conceptualização e a localização, ambas desenvolvidas num ambiente virtual, onde a comunicação, a partilha, a representação do conhecimento e a sua localização são apoiadas, muitas vezes, em ambientes digitais de comunicação, nem sempre totalmente desenvolvidos para integrar todos os processos e as necessidades inerentes a cada um deles ou à comunicação entre os especialistas e entre estes e terminólogos e/ou gestores de conhecimento.

Existe outra analogia entre o processo de localização e o de localização de ontologias quando se considera o “leitor alvo”, que, em ambos os casos, não é um leitor clássico, assumindo antes o papel de utilizador, o que obriga a que a conceptualização e a sua localização tenham em conta (1) o uso potencial que o utilizador fará do conteúdo localizado e (2) as restrições impostas pela aplicação, pelos propósitos e as orientações fornecidas pelos especialistas/clientes.

Por outro lado, o conteúdo a localizar, que assume a forma de uma representação conceptual do conhecimento do domínio, não seguirá um caminho tradicional linear, *i.e.*, do produtor de conteúdos (informação) ao seu(s) receptor(es). Pelo contrário, os intervenientes nos processos de conceptualização e localização podem desempenhar, paralelamente, o papel de produtores e receptores, o que poderá conduzir, em algumas etapas, a um processo cíclico de reconceptualização da ontologia.

Cimiano *et al.* (2010: 07), estabelecem uma outra analogia. Como descrevem, a introdução de uma aplicação num determinado país requer a adequação da sua documentação, da ajuda em linha e dos interfaces gráficos do utilizador. Este processo, que designam de *surface localization (localização à superfície)*, não afecta o núcleo da aplicação em termos do seu funcionamento. Este processo é, na opinião dos autores, similar à tradução de etiquetas no processo de localização de ontologias. No entanto, e como referem, pode dar-se o caso de as funcionalidade e o comportamento do software obrigarem a modificações, para reponderem a:

different processes and rules in place in another country. In this case we do actually change the functional core of the software in the same way we change the conceptualization to meet the requirements of a given geo-political and cultural environment.

Podemos, finalmente, estabelecer um paralelismo, que consideraremos de mais alto nível e anterior ao processo de localização, quando consideramos o processo de conceptualização de uma ontologia e o processo de internacionalização. Assim, o processo de internacionalização, tal como descrevemos antes, dimensiona um produto para que seja funcional em vários mercados, o que implica uma certa depuração das convenções culturais e linguísticas que possam perturbar este processo.

Transferindo este processo para o domínio da gestão e da representação do conhecimento, mais especificamente, para a construção, no seio de uma comunidade multilingue e multicultural, de uma representação conceptual que servirá de base à representação do conhecimento do domínio através de uma ontologia, a internacionalização pode ser percebida como um processo de generalização de um sistema de organização do conhecimento (KOS⁷⁵) tendo por base a representação dos conceitos e das relações entre estes, de modo a poder lidar com diferentes línguas e diferentes pontos de vista num ambiente organizado em rede (Boteram *et al.*, 2008: 16).

Para que esta generalização ocorra, torna-se, no entanto, necessário identificar: *“The most appropriate structural form for a core KOS and the decision for a concrete KOS that is taken as initial point for developing delocalized concepts and structures”*.

Ainda no âmbito do paralelismo entre a localização de ontologias e o processo de internacionalização, e no que respeita à produção do discurso, podemos afirmar que a relação entre produtor e tradutor já não é de um para um, uma vez que a localização requer um trabalho de equipa transversal que reúne diferentes *backgrounds* profissionais, o que, segundo Pym (2004), *“is not entirely new (...), but does accentuate the role of the mediating space, for which neither translation theory nor localization discourse has an adequate name”*.

⁷⁵ KOS – Knowledge Organization System

Por outro lado, também no caso da localização de uma ontologia, os actores envolvidos no processo de localização são, no final do processo, remetidos ao anonimato, apesar da participação activa que têm na sua construção partilhada, uma vez que, após a sua formalização, a ontologia será um elemento integrado numa aplicação que apoiará o acesso e a gestão do conhecimento do domínio, não mantendo, em princípio, qualquer referência aos participantes no processo de conceptualização e localização.

Finamente, também na construção de uma ontologia se procura criar um artefacto que, na sua génese, seja independente, em termos conceptuais, da língua em que é criado e que corresponda aos anseios de uma comunidade multilingue, i.e., que funcione num ambiente de pesquisa e gestão multilingue do conhecimento do domínio e que, tal como no processo de internacionalização, seja adequado às necessidades e especificidades linguísticas e culturais locais.

Localização de ontologias é, assim, na nossa perspectiva, o processo de adaptação linguística e cultural dos conteúdos de uma ontologia aos requisitos de uma língua e de uma cultura alvo, de modo a corresponder às suas necessidades específicas de acesso e gestão do conhecimento.

Como tal, necessita de ter em consideração, para o seu desenvolvimento, a comunicação intercultural e factores como o estatuto de uma língua, a política linguística, as preferências culturais, os usos estabelecidos e, até mesmo, as necessidades dos utilizadores, centrais ao processo de aceitação, uso e sucesso de qualquer produto, mas também a existência ou o fornecimento de serviços e tecnologias para a gestão do multilinguismo através do fluxo de informação digital a nível global.

CAPÍTULO III

3. Método para a especificação multilingue de ontologias

A consideração de aspectos multilingues no desenvolvimento de ontologias, apesar de ser uma necessidade evidente para apoiar a comunicação e gestão do conhecimento em ambientes multilingues, tem, como referimos, sido investigada de modo menos sistemático, graças à complexidade associada à construção de ontologias com recurso a mais de uma língua. Acresce, ainda, que só mais recentemente se notou uma evolução nos sistemas técnicos e abordagens teóricas disponíveis. Estes factores resultam numa carência de ontologias formalizadas que contenham, associada, informação multilingue.

Por outro lado, no desenvolvimento de ontologias tem-se dado uma importância crescente à fase da construção partilhada de representações da conceptualização do domínio, desenvolvida por grupos de especialistas, seguida do conseqüente reforço do desenvolvimento de ferramentas e de metodologias que apoiem esta fase complexa de aquisição e partilha de conhecimento.

No entanto, é nossa opinião que esta fase não mereceu ainda a devida atenção no que se refere à relação entre a construção partilhada de uma ontologia e o seu enriquecimento com elementos multilingues, sobretudo quando contextualizada no seio de uma rede colaborativa multilingue.

Neste contexto, a terminologia e as suas ligações à representação do conhecimento sugerem caminhos alternativos para o tratamento da problemática da criação e manutenção de ontologias em várias línguas, tendo por base a conjugação de contributos dos dois tipos de trabalhos – o terminológico e o da representação do conhecimento.

O recurso às metodologias propostas pela terminologia surge como um elemento essencial, no nosso ponto de vista, para a resolução de um conjunto questões que se colocam no desenrolar do processo de especificação multilingue de ontologia, das quais se destacam as relacionadas com a definição e delimitação do domínio ou subdomínio(s) a conceptualizar, a recolha, classificação e conversão de recursos semânticos mono e multilingues e o modo de integração e (re)utilização de recursos e ferramentas linguísticas já disponíveis.

Por outro lado, a construção de estruturas conceptuais representativas do conhecimento do domínio é, há muito, prática corrente no trabalho e na pesquisa terminológica, pelo que, os modelos terminológicos referentes à representação do conhecimento do domínio são uteis para explorar a possibilidade de transformar aquelas estruturas em ontologias. Acreditámos também, em concordância com Budin (2005), que o recurso à terminologia contribui para capturar de forma mais profunda a relatividade intercultural e muitas das assimetrias existentes entre conceptualizações específicas às diferentes culturas presentes numa comunidade.

Tendo em conta este contexto e os princípios enunciados, bem como as lacunas identificadas ao longo da análise do estado da arte, desenvolvemos a ideia de construir um método de apoio à especificação multilingue de ontologias na sua fase de conceptualização, com recurso à localização.

O recurso à localização surge porque, como explicitámos já, é nossa opinião que a localização não é apenas uma questão de língua, mas também de conhecimento. A localização preocupa-se também com o conhecimento sobre o discurso (termos complexos, colocações terminológicas) e permite, se bem executada, a correcta gestão e transferência da informação e do conhecimento, nomeadamente enquanto etapa integrante da gestão do conhecimento baseado em ontologias.

Este método foi desenvolvido com o fim de, por um lado, contribuir para a resolução de parte dos problemas e lacunas identificados – de que se destacam a dificuldade inerente à escolha de metodologias a adoptar para a representação do conhecimento num ambiente de conceptualização e a localização de ontologias e o número diminuto de ferramentas e ambientes específicos que apoiem este processo - e, por outro, auxiliar aos processos de construção partilhada de uma ontologia e da sua localização, tendo em conta a participação dos especialistas de domínio.

A construção do método seguiu uma perspectiva que tem como base uma metodologia de análise interlinguística de suporte à conceptualização do domínio. Para o seu desenvolvimento assumimos que o processo de conceptualização é o ponto de partida para o desenvolvimento da representação do conhecimento do domínio em mais do que uma língua natural. Tem como elementos centrais ao seu desenvolvimento e aplicação o uso de metodologias de base terminológica, num enquadramento teórico que privilegia uma

abordagem conceptual à representação do conhecimento e à sua localização, mas que considera os recursos e contributos resultantes da abordagem linguística.

3.1. Objectivos e justificação do método

O método tem como objectivo específico apoiar especialistas e terminólogos no desenvolvimento de uma especificação multilingue de uma ontologia. Propõe, para tal, um espaço e um fluxo de trabalho colaborativo com o objectivo de:

1. apoiar a localização de ontologias para diferentes línguas no momento da sua conceptualização, a partir das representações conceptuais do domínio,
2. permitir a integração dos contributos resultantes da interacção entre os diferentes interlocutores envolvidos na conceptualização, bem como a eventual reconceptualização das estruturas conceptuais,
3. fomentar o processo de partilha do conhecimento do domínio no contexto da rede colaborativa multilingue,
4. permitir que se considere, no momento da conceptualização, a eventual necessidade de adequação da ontologia aos diferentes universos linguísticos e culturais a que se destina,
5. facilitar a evolução e manutenção da ontologia ao diminuir a necessidade de reconceptualização após a sua especificação formal.

Importa salientar que o método que propomos não foi desenvolvido tendo como fim a recolha e construção prévia de uma terminologia para cada uma das línguas dos parceiros da rede ou de uma terminologia multilingue de domínio. Também não é nossa intenção utilizar uma ferramenta que localize de modo (semi)automático uma ontologia previamente formalizada.

Na nossa proposta, a localização acontece no momento seguinte ao de uma conceptualização inicial e decorre no espaço de conceptualização, sendo a representação multilingue dos seus resultados desenvolvida e disponibilizada aos especialistas com recurso a mapas conceptuais. Não se parte, ao contrário do que acontece noutras metodologias (ex.:

NEON, 2010), da ontologia formalizada, mas, sim, da organização semiformal do conhecimento, disponibilizada sob a forma de mapas conceptuais.

De facto, as abordagens mais correntes à localização de ontologias, tal como descrito no capítulo anterior, são desenvolvidas recorrendo:

- a uma língua comum, normalmente o Inglês, que é depois localizada de acordo com as necessidades e os propósitos da rede, ou
- às línguas dos diferentes especialistas que colaboram numa rede, sendo depois feito um mapeamento (semi)automático, a partir, por exemplo, de métodos estatísticos, entre os conceitos das diferentes línguas.

Num caso e noutro, a localização acontece após a especificação formal da ontologia, tal como se representa em seguida:

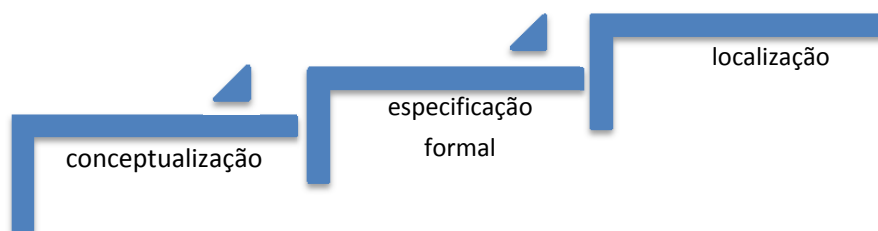


Fig. 7 - Abordagem típica à localização de ontologias

A localização é, assim, vista sobretudo como um problema *a posteriori*⁷⁶. Um exemplo já abordado é o do recurso a ferramentas como o *LabelTranslator*, uma ferramenta usada para localizar uma ontologia já construída e formalizada, e não um elemento auxiliar à sua construção inicial, ainda que ofereça um fluxo de trabalho tendente a facilitar a eventual revisão posterior da ontologia. Segundo os autores, a aplicação usa como *input* uma ontologia ou um conjunto de ontologias cujas etiquetas (termos) estão especificadas numa

⁷⁶ Na nossa opinião, esta abordagem não corresponde na totalidade às necessidades de uma rede colaborativa, onde a necessidade de informação localizada surge mais cedo, em função do ciclo de vida curto que é expectável que a rede tenha ou das necessidades de obtenção de uma resposta célere a oportunidades de negócio. De facto, esta limitação temporal pode obrigar à disponibilização dos elementos multilingues que apoiem a comunicação especializada e a gestão e representação da informação ainda na fase de conceptualização do domínio.

dada língua natural e obtém a tradução *mais provável* para cada etiqueta (termo) da ontologia numa língua natural alvo.⁷⁷

Pretendemos, assim, afastarmo-nos da localização dos termos, entendidos apenas como meras etiquetas de um sistema, ao criar as condições para que o processo de localização possibilite a escolha dos equivalentes apropriados tendo por base um ponto de vista conceptual.

Neste sentido, procurámos criar as condições para que, durante o processo de localização, o enfoque não recaísse apenas em encontrar o equivalente mais correcto do ponto de vista da língua, a partir, por exemplo, do uso de aplicações automática ou semiautomáticas. Procurámos antes, envolver activamente os especialistas de domínio no processo de localização, como forma de considerar os diferentes universos linguísticos e culturais e obter o contributo daqueles na escolha do equivalente mais correcto.

Daí que advoguemos, como expressámos antes, a necessidade de associar o trabalho terminológico ao de localização e de representação do conhecimento, seguindo os princípios e práticas terminológicas e tendo como ponto de partida uma abordagem onomasiológica, que parte do conceito, não se concentrando a localização, assim, na tradução dos nomes, mas, antes, na procura das denominações naturais que, em cada língua, correspondem a um conceito especializado.

Para esse efeito, o método que a seguir apresentamos possui um cariz integrador, i.e., parte de uma base terminológica, mas considera, no seu desenvolvimento e aplicação, elementos de outras disciplinas, nomeadamente, da localização, da tradução e da representação do conhecimento.

3.2. Visão geral do método

Os conceitos descritos pelas ontologias não se restringem às palavras ou termos em uso numa língua natural particular, mas, como expusemos antes, a prática conduziu a que se construíssem as ontologias sobretudo em língua inglesa. Esta prática tornou-se redutora

⁷⁷ <http://sinai.ujaen.es/timm/wiki/index.php/LabelTranslator>

para os objectivos a alcançar com o uso de uma ontologia, nomeadamente, o de permitir a interoperabilidade semântica em ambientes que disponibilizem um grande número de recursos heterogéneos, e o de potenciar a extracção e o acesso à informação e ao conhecimento existente nas várias línguas.

Tendo em conta este contexto, procurámos desenvolver um método que enriqueça a fase de conceptualização⁷⁸ ao dotar os seus interlocutores - conjunto de parceiros de uma rede colaborativa provenientes das áreas científica e técnica de um domínio específico - da possibilidade de acederem à representação conceptual e de a localizarem, o que viabiliza a análise da representação conceptual também na presença da sua língua de origem, e não só na língua de trabalho convencionada para a comunicação na rede.

Para a construção da nossa proposta assumimos, assim, que, no caso da conceptualização e da localização, os dois processos devem ocorrer consecutivamente, tal como se representa na figura seguinte, uma vez que tal possibilita que numa comunidade se considerem as diversas perspectivas e a informação disponível das diferentes línguas na conceptualização do domínio.

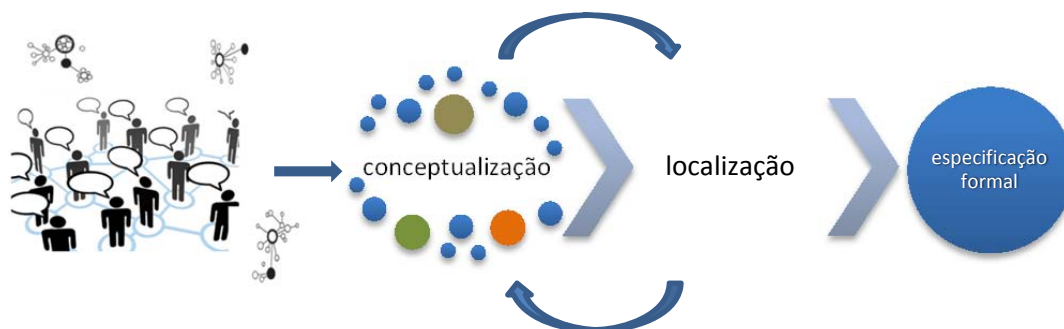


Fig. 8 – Abordagem proposta à localização de ontologias

⁷⁸ A conceptualização situa-se, como descrevemos antes, na fase inicial da construção da ontologia, fase que Pereira & Soares (2008: 608) designam de *especificação semiformal*, em que não se recorreu ainda a uma representação formal. As representações socorrem-se, nesta fase, das mais variadas fontes de análise, nomeadamente das textuais, taxonómicas e de outras representações conceptuais, sendo alvo de constante negociação através da colaboração directa dos especialistas de domínio. O objectivo é conduzir à co-construção de recursos semânticos (Pereira & Soares, 2008: 608), recursos que, no nosso caso, assumirão uma dimensão multilingue.

Para tal, adoptámos na nossa abordagem um ponto de vista ligado aos processos de gestão terminológica, acrescido dos da localização e da gestão do conhecimento, de modo a possibilitar a localização de um dado termo da ontologia em adequação ao universo linguístico e cultural de uma determinada língua e a capturar as especificidades de tradução entre as diversas línguas.

Na nossa opinião este é um processo dinâmico em que os termos podem desempenhar um papel importante no acesso e na partilha do conhecimento especializado e em que a presença de mais do que uma língua tenderá a potenciar a partilha do conhecimento entre os membros da comunidade.

O carácter iterativo e, até certo ponto, cíclico, que existirá entre o desenrolar dos processos de conceptualização e localização pretende possibilitar, também, uma primeira partilha de conhecimento mais rica e uma análise mais aprofundada da informação disponível, bem como a eventual reconceptualização da representação conceptual. Tal acontece dado que os diferentes especialistas têm a possibilidade de considerar e disseminar diferentes pontos de vista, não se concentrando ou questionando apenas sobre a representação após a sua formalização, como tipicamente ocorre nas abordagens antes descritas (ponto 2.4).

O uso de um processo consecutivo de conceptualização e localização poderá, deste modo, concorrer para o acesso mais imediato a um conjunto mais vasto de conhecimento sobre o domínio e permitir aos especialistas uma negociação e estruturação do conhecimento do domínio mais completa e mais representativa e, como consequência, uma especificação mais granular da(s) ontologia(s), adequada às diferentes realidades linguísticas. Constitui-se, por outro lado, como um factor de intercompreensão capaz de auxiliar no processo de negociação, ao diminuir o grau de ambiguidade e ao contribuir para a eliminação de elementos que possam concorrer para a confusão conceptual e terminológica.

O método que propomos procura, assim, evitar a situação descrita por Declerck *et al.* (2010), para os quais a adaptação de uma conceptualização⁷⁹, - que consideram ser necessária em casos nos quais a ontologia ou a conceptualização original não reflecta a organização de certos aspectos da semântica do domínio, tal como é expressa na língua de

⁷⁹ Que designamos de *reconceptualização*.

chegada, - requer uma *inspecção manual*, a qual não está incluída nas actividades a desenvolver no processo de localização. Os autores prevêem apenas a detecção de eventuais conflitos de tradução a serem resolvidos *a posteriori*, através da identificação de problemas de tradução que apontariam para eventuais adaptações a serem resolvidas manualmente, devendo a actualização ser levada a cabo num processo separado.

Por outro lado, a abordagem proposta incorpora e segue princípios já em uso nos modelos adoptados na localização de produtos digitais, nomeadamente o modelo “sim-ship”⁸⁰, o qual prevê a produção simultânea de versões localizadas da versão original. Este modelo, adequado a este contexto, tem implicações na estabilidade do trabalho de localização, uma vez que há que ter em conta e gerir a dinâmica de evolução a que o sistema conceptual pode estar sujeito.

No entanto, no contexto de um espaço de uma conceptualização existem, na nossa perspectiva, factores diferenciadores que atenuam a possível instabilidade criada pela simultaneidade do processo, nomeadamente a presença dos especialistas, uma delimitação bem definida do domínio de especialidade em análise e uma noção concreta e aceite pelos vários interlocutores quer dos propósitos e requisitos que orientam a construção da ontologia quer dos objectivos para o seu uso posterior.

Este método pretende, por outro lado, criar as condições para estruturar o processo de localização, através da identificação de um fluxo de trabalho e de um conjunto de ferramentas de auxílio à tradução, como bases de dados terminológicas, lexicais, entre outras, que assistam os especialistas e restantes interlocutores no desenvolvimento do processo.

Para tal, será disponibilizado aos especialistas um espaço virtual de conceptualização e de localização, que permita sustentar a situação comunicativa específica e que disponibilize ferramentas de trabalho que apoiem o acesso ao conhecimento e à sua representação em termos multilingues, através do recurso a mapas conceptuais, acessíveis de forma partilhada, nos quais se representem os conceitos do domínio e as relações que

⁸⁰ Simultaneous shipment

estabelecem entre si, bem como os seus equivalentes nas línguas dos actores da rede colaborativa.⁸¹

O método proposto procura, deste modo, ser integrador e criar um fluxo de trabalho que conduza a um enriquecimento linguístico gradual das representações conceptuais, e permita a interacção constante entre especialistas e recursos lexicais, terminológicos, entre outros; a interacção com aplicações de tradução e localização; bem como a interacção entre os especialistas, como se apresenta na imagem seguinte.

O método tem, por outro lado, em conta as diferentes fases para a construção de uma ontologia representada em várias línguas, mas assume como necessária a existência de uma representação conceptual inicial, como base para o enriquecimento multilingue dessa representação - partimos de uma representação partilhada e consensual do conhecimento do domínio para determinar o modo como é expresso nas várias línguas.

⁸¹ No fundo, pretende-se proporcionar um espaço que, através da inclusão e integração dos meios, e não da sua multiplicação, potencie as dimensões expressivas/comunicativas mais complexas, projectando um espaço de autonomia e de partilha de conhecimento; espaço construído a partir de uma abordagem necessariamente transdisciplinar que considere o acesso e a partilha de conhecimento, a sua representação e que garanta a sua especificação formal.

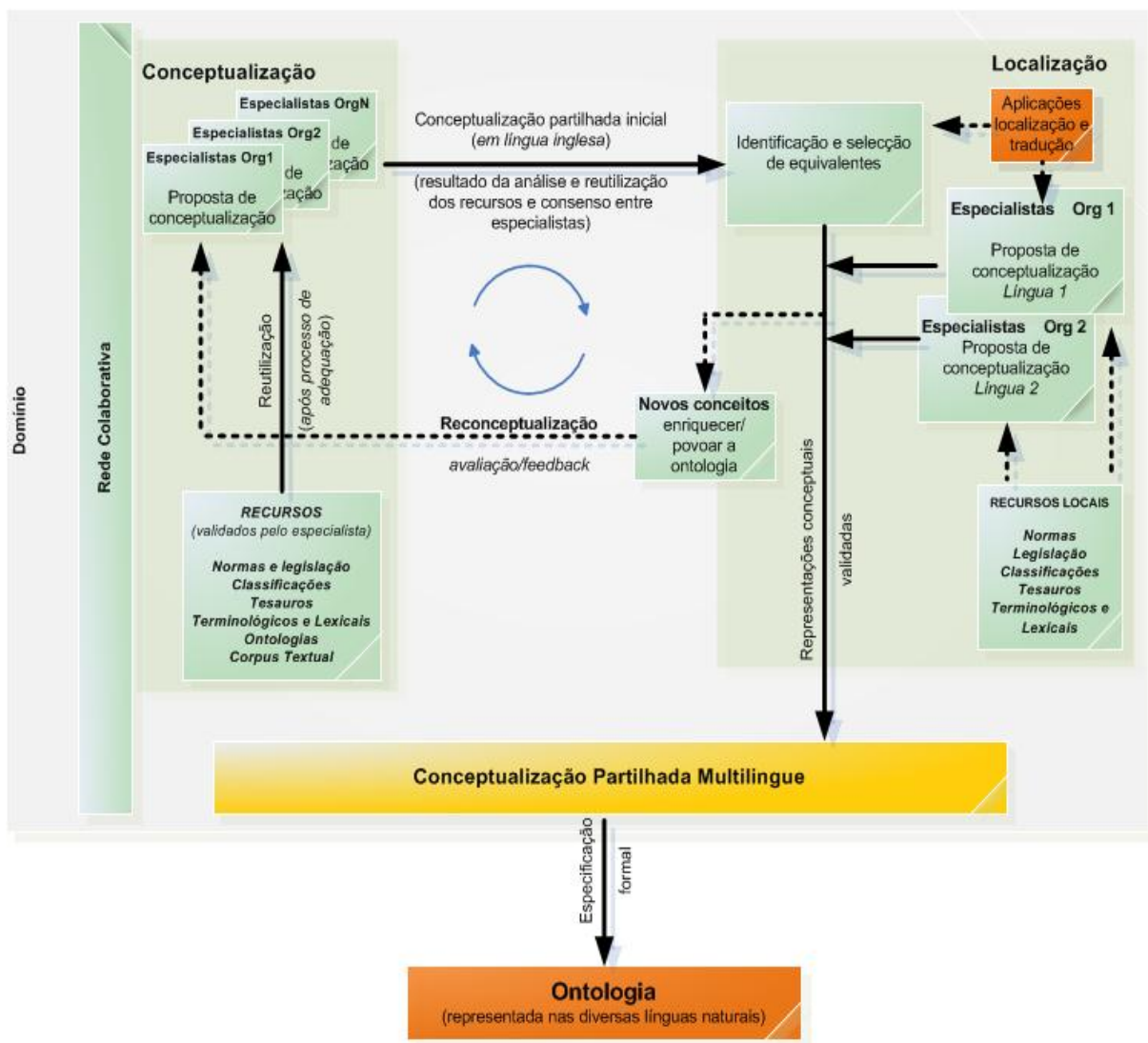


Fig. 9 - Método para a especificação multilingue de ontologias

Visão de alto nível

Acreditamos que o carácter recursivo e dinâmico proposto pelo método acima apresentado se adequa ao processo de representação do conhecimento de um domínio no contexto de uma rede colaborativa. De facto, a presença dos termos nas várias línguas pode servir de incentivo a um conjunto de operações conceptuais, como o acesso a conhecimento novo, que pode ainda não ter sido expresso na língua do especialista, ou o reequacionar da representação que o próprio faz do domínio, quer partindo da sua visão do mundo quer da necessidade de a adequar e reestruturar, tendo em conta os contributos e a visão de outros especialistas e de outras culturas.

Importa, no entanto, explicitar que a pretensão deste método é, em primeiro lugar, conduzir à construção de uma ontologia representada em mais do que uma língua, através da sua localização, podendo este processo implicar ou não um processo de reconceptualização.

Sendo a representação conceptual expressa por via da língua através dos termos, consideramos que, se a conceptualização inicial for partilhada e validada por todas os especialistas implicados na actividade de localização, o processo de localização centrar-se-á apenas na escolha dos equivalentes. Só no caso da existência de lacunas, da emergência de conhecimento novo ou da falta de consenso é que o especialista poderá propor uma reconceptualização.

A especificação de um método que contribua para que a localização possa potenciar a interoperabilidade semântica entre recursos de conhecimento heterogéneos no âmbito de um domínio específico requer que tenhamos em conta o contexto de estruturação e de uso do conhecimento e os meios disponíveis para a apreensão da complexidade e riqueza semântica do domínio e para a sua representação em termos multilingues. Exige também, como referimos antes, que consideremos os objectivos e os propósitos de uso desse conhecimento pela comunidade de utilizadores potenciais.

Nesse sentido, e tal como sistematizámos na visão de alto nível, acima representada, serão disponibilizados, no âmbito da aplicação do método, diferentes tipos de recursos para apoiar o processo de conceptualização partilhada do domínio e o processo de localização, recursos que descrevemos em seguida.

3.3. Contributo da terminologia

3.3.1. Constituição do *corpus*

Como ressalta da representação do método proposto (ver fig. 10), o primeiro elemento a considerar - e que assumimos como central a todo o processo - é o domínio de especialidade. O domínio ou subdomínios a serem representados são, na nossa opinião, determinantes para o desenrolar do processo de conceptualização e de localização, sendo

que a sua conceptualização depende, em grande parte, dos propósitos e das actividades a desenvolver, condicionando, assim, quer o trabalho terminológico quer o ontológico.

Tal como postula a norma ISO 704:2009(E) (2009: 08):

In organizing concepts into a concept system, it is necessary to bear in mind the subject field that gave rise to the concept and to consider the expectations and objectives of the target users. The subject field shall act as the framework within which the concept field, the set of thematically related but unstructured concepts, is established⁸².

O desafio coloca-se, assim, em partir da multiplicidade de contributos dos especialistas do domínio ou do contributo da diversidade de textos e de outros recursos semânticos para a construção de uma ontologia multilingue axiomatizada e formalizada.

Assim, e tendo em conta a abordagem de base terminológica que propomos para a implementação do método, um dos passos essenciais para apoiar o desenvolvimento dos processos de conceptualização e localização consiste na constituição de um *corpus* específico do domínio, composto por recursos representativos das diferentes línguas. Estes devem ser recolhidos tendo em conta as necessidades e propósitos previamente definidos, de modo a serem utilizados como ferramenta de apoio à representação do conhecimento do domínio, ao permitir, por exemplo, a extracção terminológica ou o acesso aos contextos de uso dos termos.

O trabalho terminológico, ainda que não sendo o único meio de auxílio à representação do conhecimento especializado, assume neste processo um papel essencial e estruturante. A língua de especialidade é o instrumento básico de comunicação entre os especialistas, surgindo a terminologia, com naturalidade, como o elemento mais relevante no processo de explicitar cognitivamente o sistema de denominação. De facto, a terminologia, além de ordenar o seu pensamento, é essencial ao especialista no momento de estruturação e disseminação do seu conhecimento sobre o domínio, em uma ou mais línguas⁸³.

⁸² Nosso sublinhado.

⁸³ Nas palavras de Cabré (2004), a ordenação do pensamento e a conceptualização representam a dimensão cognitiva da terminologia, ao passo que a transferência de conhecimento constitui a sua dimensão comunicativa.

No entanto, partir de um *corpus* para o trabalho de representação do conhecimento de um domínio, coloca, na nossa perspectiva, limitações, como referimos no ponto 1.4.1. e 1.4.2. De facto, nos textos de especialidade são denominados conceitos e são estabelecidas relações semânticas que denotam parcialmente as relações conceptuais que resultam do processo cognitivo. O texto é, ele mesmo, um sistema complexo que, contudo, não se sobrepõe à complexidade do conhecimento representado (Costa, 2010: 03). Entender estas limitações torna-se, assim, essencial ao bom uso dos textos de especialidade em processos de conceptualização.

3.3.1.1. Recursos linguísticos e semânticos

O processo de constituição de um *corpus* surge normalmente associado à constituição de um *corpus* textual. No entanto, o *corpus* textual surge cada vez mais acompanhado de outros tipos de recursos, cuja utilização tem ganho uma maior atenção e relevância, sobretudo à medida que o paradigma da reutilização de recursos se vai afirmando como necessário e útil, dada a escassez de tempo e a dificuldade e custos que comporta construir de raiz recursos para a representação do conhecimento de um domínio específico ou para a sua localização ou tradução. Esses recursos são do tipo⁸⁴:

1. Lexicográficos e terminológicos, como:
 - a. Dicionários
 - b. Terminologias
 - c. Glossários
 - d. Bases de dados terminológicas
2. Representações conceptuais, como:
 - a. Classificações
 - b. Tesouros
 - c. Taxonomias
 - d. Ontologias (de alto nível, de domínio, ...)
 - e. Gráficos (mapas de conceitos, esquemas, diagramas de UML, ...)

⁸⁴ Não existe uma classificação definida ou de referência sobre o conjunto dos elementos que podem ser considerados na definição de uma tipologia de recursos semânticos. Apresentamos, no Anexo 1, uma sistematização das propostas mais utilizadas.

Cada um destes recursos possui, como características que podemos considerar comuns, um conjunto de padrões que permitem:

- 1) que sejam considerados independentemente da sua origem;
- 2) que se possam estabelecer parâmetros para a sua avaliação, validação e uso;
- 3) que se possa aceder e considerar o modo como representam o conhecimento;
- 4) que possam ser reutilizados.

Estes recursos são de grande utilidade não só para o terminólogo - que os utiliza como um meio para melhor e mais rapidamente aceder a representações do conhecimento já estruturadas - mas também para o especialista, que os pode utilizar como um ponto de partida para as actividades de delimitação do seu domínio ou subdomínio e ao longo de todo o processo de conceptualização.

O uso destes recursos, sobretudo dos terminológicos e de classificações, taxonomias, tesouros e ontologias, leva a que parte da informação relevante para os especialistas surja, assim, já sistematizada em artefactos de representação do conhecimento do domínio. Estes disponibilizam, de modo relativamente uniforme e acessível, informação originária de diferentes fontes e, ainda que menos frequentemente, de diferentes línguas, uma vez que um número significativo destes recursos resulta do trabalho de harmonização e normalização desenvolvido por organismos internacionais ou do trabalho desenvolvido no âmbito de projectos multinacionais.

Consideramos, também, que estes recursos podem contribuir para um melhor entendimento entre o especialista e o terminólogo, quando ambos participam activamente na conceptualização⁸⁵, ao, por exemplo, apoiarem o processo de análise e definição dos conceitos centrais ao domínio ou ao apoiarem uma identificação dos pontos de contacto com outros domínios importantes para a compreensão do conhecimento do domínio em análise.

⁸⁵ Costa (2010: 04) distingue entre terminólogo linguista e terminólogo engenheiro, afirmando, a este propósito, que: « *Il est vrai que, par sa formation, le terminologue linguiste a tendance à privilégier les corpus textuels, alors que le terminologue ingénieur comprend mieux la schématisation des connaissances* ». A autora retrata o uso destes elementos tendo em conta as abordagens passíveis de serem adoptadas pelos terminólogos: « *Le terminologue ingénieur dira que les schémas sont plus proches des connaissances, parce que la représentation est plus formelle donc moins ambiguë. Et puis surtout, ils n'ont pas à gérer l'ambiguïté des discours. Quant au terminologue linguiste, il préfère gérer les discours malgré leurs ambiguïtés, car de par sa formation il a les outils nécessaires à la réduction, voire à l'annulation de ces ambiguïtés* ».

Para além disso, e como atesta Budin (2007b), o uso destes recursos na construção de ontologias permite uma exploração cabal de todos os tipos de:

languages resources (LR) and knowledge organization systems (KOS), providing a framework not only for their semantic enrichment and formalization as ontologies but also for ontology-based multilingual authoring, text generation and translation.

De modo a obter um conjunto relevante e representativo, a selecção destes recursos semânticos deve, na nossa opinião, atender a critérios de avaliação como a sua origem, acessibilidade, estado de desenvolvimento, granularidade ou relevância, bem como outros a determinar, uma vez que cada recurso foi concebido por um conjunto de especialistas num determinado momento, de acordo com uma perspectiva e uma necessidade de organização do conhecimento, tendo sido estruturado com propósitos, princípios e convenções próprias, cada uma baseada no seu próprio modelo prototípico de dados, a que acrescem ainda algumas formas híbridas que estes recursos vão assumindo à medida que a necessidade de acesso e gestão do conhecimento especializado aumenta.

Propomos, em seguida, como elemento de apoio à aplicação do método, uma sistematização dos critérios passíveis de serem aplicados por terminólogos e especialistas para as actividades de selecção destes recursos e para a sua validação. Estes critérios não pretendem ser exaustivos e deverão, na nossa perspectiva, ser adequados em função da maior ou menor especificidade das actividades a desenvolver e do conhecimento a representar, bem como ter em conta os diferentes tipos de utilizadores/especialistas.

Tipo de recurso	Critérios	Descrição	Questões a considerar
	Origem	Autores	Tipo de organismo (Instituto de I&D, Organismo de normalização, Consórcio de projecto da UE, etc.)?
	Relevância	Para o domínio e subdomínio; para os objectivos do projecto	Ex.: Tamanho do <i>corpus</i> , especificidade, etc.?
	Adequação	Do ponto de vista do especialista do domínio Do ponto de vista do gestor do conhecimento	Adequa-se às necessidades e propósitos do projecto?
	Granularidade	Domínios ou subdomínios abrangidos	Geral? Específico?

Terminologias Ontologias Classificações Taxonomias Tesauros Base dados	(Não)/Normalizado	Implementado como norma Outro tipo	Uso generalizado? Uso específico?
	Facilidade de uso	Facilidade com que os utilizadores acedem e interagem com o recurso	Representado de modo convencional? Existe informação sobre o seu uso?
	Interoperabilidade	Linguagem em que está disponível Possibilidade de acesso e reutilização	Formato: Owl, rdf,...? Acessível num repositório?
	Línguas naturais	Características multilingues	Possui ou suporta características multilingues?
	Princípios de design	Estrutura interna	Tipo de conceitos: identifica ou descreve os conceitos genéricos e as relações?
	Documentação disponível	Tipo de documentação Disponibilidade Acessibilidade	Possui documentação de Apoio? Disponibiliza informação sobre os pressupostos, propósitos, etc. que presidiram ao seu desenvolvimento?
	Uso Prévio	Utilização em outros projectos	Resultados obtidos?
	Estado actual	Estado de desenvolvimento	Terminado, em desenvolvimento, em revisão?

Tabela 1 – Critérios para a selecção de fontes⁸⁶

Este processo de selecção acontece sob influência de um conjunto de condicionantes que podem incluir, por exemplo, os constrangimentos em termos temporais⁸⁷ e de recursos de conhecimento⁸⁸; a necessidade de adequação das diferentes tipologias e formatos destes últimos ao seu uso no processo de construção da ontologia; as ontologias a reutilizar e a sua semântica; os propósitos da integração ou mapeamento destas; os utilizadores potenciais e,

⁸⁶ Os critérios definidos nesta tabela constituem uma actualização da proposta apresentada na Conferência TKE 2005 - Silva, Manuel, Simões, D., Costa, R. e Soares, A. (2005). *Know-Construct: Knowledge based community management in the Construction Industry*. In Proceedings of the 7th Terminology and Knowledge Engineering Conference, Workshop Working with Terminology and Knowledge Management Systems. Copenhaga: GTW., tendo sido desenvolvidos no âmbito do projecto europeu Know-Construct (ver Apêndice A).

⁸⁷ A dinâmica imposta à conceptualização e especificação da ontologia pode levar a que não se considerem todos os recursos disponíveis.

⁸⁸ A inexistência destes pode condicionar os resultados da qualidade da conceptualização e da localização.

num momento posterior, o *feedback* obtido a partir da aplicação do resultado num sistema de gestão do conhecimento de um domínio.

Assim, depois de classificados e validados, estes recursos, poderão ser usados, por parte do terminólogo, do(s) especialista(s) ou do gestor de conhecimento, como elementos de apoio à análise e representação do conhecimento do domínio, nas diferentes fases do processo de conceptualização e validação do sistema conceptual, podendo funcionar como *input* para o desenvolvimento de representações de conhecimento mais avançadas (específicas) e para a partilha e integração de conhecimento novo (ainda não representado).

Uma vez seleccionados, a heterogeneidade dos recursos implica a definição de uma metodologia que permita a sua reutilização e integração⁸⁹. De facto, a análise e observação destes recursos podem fornecer dados importantes para a conceptualização do domínio, embora não deva, na nossa opinião, ser baseada em processos totalmente automáticos de transposição.

É nossa percepção que a reutilização dos conceitos resultantes destes recursos influencia a conceptualização⁹⁰ e a localização, pelo que deve ser realizada de modo ponderado e com a intervenção de um terminólogo, na medida em que:

- a localização vai incidir sobre o resultado da adequação de um determinado recurso, construído de acordo com princípios e propósitos específicos, para outro, novo e de tipo e propósitos diferentes,
- a conjugação e a interligação entre os diferentes recursos pode, sobretudo no caso dos recursos já representados em diferentes línguas, influenciar a selecção dos equivalentes⁹¹.

⁸⁹ Um exemplo advém do uso de recursos do tipo lexicográfico, que, como descreve Vieu (2011), são, tipicamente: “*much larger and likely to be built by lexicographers and terminologists rather than formal ontologists, or even to be semi-automatically built from other resources and data*”, o que obriga a precauções na sua adequação e reutilização.

⁹⁰ Como explicita Roche (2006), a propósito dos problemas levantados pela construção de ontologias, quer em termos linguísticos quer em termos conceptuais, “*The lexicon of a LSP (...) and the vocabulary of a concept’s names are in fact two different things which should be study independently, and then connected if necessary*”. Para o autor, com o qual concordamos, o léxico de uma língua de especialidade (LSP) é composto de palavras de uso que podem ser polissémicas, ao passo que a designação do conceito é, por definição, monossémica.

⁹¹ No que concerne a este ponto, e de modo a evitar uma escolha incorrecta de equivalentes, o terminólogo deve, em cooperação com os especialistas e com recurso ao *corpus* textual, verificar a adequação dos termos a partir do seu uso em diferentes situações em discurso. O *corpus* serve assim uma dupla função: a de

A (re)utilização destes recursos deve, assim, ter em conta um conjunto de critérios e cuidados, sendo necessário ter em atenção:

1. até que ponto são reutilizáveis para efeitos da conceptualização do domínio;
2. os cuidados a ter no mapeamento semântico entre termos de dois ou mais recursos no contexto multilingue⁹²;
3. que elementos desses recursos poderão estar ou estão disponíveis para reutilização (conceitos, relações, definições,...);
4. o modo como poderão ser adaptados às necessidades predefinidas e aos propósitos da conceptualização e da localização;
5. a sua diversidade, quer no que respeita aos princípios teóricos que presidiram à sua definição quer no que se refere ao *design* adoptado para o seu desenvolvimento⁹³.

De modo a possibilitar e a facilitar o processo de análise, uso e (re)estruturação do conhecimento disponível nos diferentes recursos semânticos e a comunicação entre os diferentes interlocutores, propomos que estes sejam, dada a sua heterogeneidade, transformados em mapas conceptuais⁹⁴.

permitir a identificação e extracção dos termos do domínio e a de verificar e validar o uso dos equivalentes seleccionados.

⁹² Esta preocupação coloca-se-nos quando procuramos reutilizar e harmonizar, num mesmo mapa conceptual, o uso de termos provenientes, por exemplo, de classificações e tesouros. Neste caso, estamos perante diferentes tipos e níveis de informação que ocupam posições diferentes no recurso de onde são originários, factos que têm que ser considerados para a correcta representação do conhecimento do domínio e para a construção da ontologia.

⁹³ Hubrich *et al.* (2008) identificaram, na sua proposta para a construção de ontologias em mais do que uma língua, outros problemas e restrições à reutilização de sistemas de organização do conhecimento, de que destacamos, no processo de reutilização, o facto das relações conceptuais do sistema de organização do conhecimento nem sempre serem tidas em conta. *“Consequently the newly created KOS are useful to improve recall and precision of retrieval processes but they are not suitable for cross-national browsing and query modification purposes as the semantic relations of the connected KOS are not compatible”*.

⁹⁴ A proposta de uso de mapas conceptuais segue uma abordagem consensual na comunidade que se dedica à representação e gestão do conhecimento. O seu valor para a representação do conhecimento será analisada no ponto 3.6. A integração do conhecimento proveniente destes recursos pode assumir duas formas: (1) integrar todo o conhecimento numa única estrutura, o que, sobretudo numa estrutura de alto nível, se revela uma impossibilidade, dado o elevado número de fontes e modos pré-existentes de estruturar, representar e perceber o conhecimento do domínio; (2) considerar as diferentes representações de conhecimento e os diferentes recursos como parte de um sistema mais lato e construir o sistema conceptual por módulos interdependentes que, no seu conjunto, venham a representar o conhecimento de domínio e permitam a sua evolução e a acomodação do conhecimento novo que possa surgir. A nossa opção é claramente pela segunda abordagem, uma vez que permite, na nossa opinião, uma análise mais cuidada de cada módulo.

Para tal ter-se-ão que conjugar as competências do terminólogo - advindas da sua experiência da gestão de terminologia e da organização e representação do conhecimento de um dado domínio - com as competências dos especialistas, advindas do seu conhecimento do domínio e da sua capacidade para o representar através de um sistema conceptual.

Nesta etapa, a terminologia e o terminólogo representam um papel fundamental na adequação e transposição da informação terminológica e conceptual entre os diferentes tipos de recursos de conhecimento disponíveis e no auxílio à estruturação e representação dos conceitos provenientes de diferentes fontes, bem como no processo da sua avaliação e validação por parte dos especialistas.

Por outro lado, o terminólogo pode auxiliar na promoção da comunicação entre especialistas, de modo a que explicitem as decisões e identifiquem falhas na organização conceptual, sobretudo ao nível dos conceitos não representados. O terminólogo pode, ainda, agir de modo a que os recursos sejam apresentados e analisados por etapas, de acordo, por exemplo, com a sua relevância e abrangência, de forma a não haver um excesso de elementos em análise em simultâneo.

3.3.1.2. Corpus Textual

Quando construímos um sistema conceptual ou uma ontologia, o recurso à língua natural é inevitável, uma vez que é através dela que manifestamos verbalmente as definições dos conceitos, por via dos termos. O acesso ao discurso do especialista torna-se, assim, um elemento importante, sendo através do texto de especialidade que o especialista contribui de forma decisiva para a evolução do conhecimento; é a ele que o especialista recorre para transmitir e aceder a esse mesmo conhecimento.

Revisitando as palavras de Costa (2010: 02) o saber que o indivíduo detém sobre uma ciência consiste, na essência, *“no conhecimento e na capacidade de criação de discursos e de textos que poderão servir de referência a outros ou a futuros membros da comunidade científica”*. No entanto, e como deixa claro a autora, o texto também é o espaço do debate, conseqüentemente, *“é o espaço do não consenso entre os membros de uma comunidade de*

especialistas, é o espaço da construção da polissemia, da ambiguidade, da sinonímia, da imprecisão voluntária ou involuntária” (Costa, 2005: 09).

A autora vai mais além e aponta grandes dificuldades no recurso ao texto para o desenvolvimento de um trabalho terminológico de base conceptual e para a representação do conhecimento, tal como já referimos no ponto 1.4.1. No entanto, e apesar das dificuldades reconhecidas e dos cuidados necessários na análise dos resultados obtidos quando recorremos ao texto e extraímos informação para efeitos de conceptualização de um domínio ou para a construção de ontologias, consideramos a existência de um *corpus* textual como um meio que permite a obtenção de resultados interessantes⁹⁵ capazes de auxiliar no processo de organização e representação do conhecimento de um domínio específico. O recurso ao *corpus* textual é, na nossa perspectiva, importante sobretudo se usado como parte de abordagem mista, que considere a primazia da perspectiva conceptual mas que complemente esta com a linguística, tal como explicitámos no ponto 1.4.2.

O recurso a um *corpus* textual⁹⁶ que contribua para a identificação dos candidatos a termos ou de contextos ricos em informação, pode tornar-se um meio importante para o desenvolvimento do processo de conceptualização do domínio em estudo. Como afirma Budin (2006), a existência de um *corpus* contribui para que se proceda a uma análise conceptual, ao explicitar a dimensão semântica, i.e. o significado dos termos. A existência de um *corpus* pode contribuir ainda noutros aspectos da construção e representação do conhecimento, nomeadamente:

1. *Help negotiate a common understanding of terms in transdisciplinary and transcultural discourse,*
2. *Reduce unnecessary synonyms, disambiguate polysems, help separate homonyms,*
3. *Support knowledge sharing and knowledge transfer in cooperative work environments,*

⁹⁵ Esta perspectiva é partilhada por diferentes autores como Lino (1996), Wright (1997), Cabré (1999, 2004, 2005), Budin (2004, 2007), Temmerman (2001, 2003), Condamines (2007) ou Despres e Szulman (2008).

⁹⁶ Na perspectiva de Lino (1996: 30), um *corpus* textual serve de base a um sem número de pesquisas e pode prosseguir uma variedade de objectivos, como a selecção e observação do comportamento de unidades terminológicas ou a selecção de contextos, entre outros. A identificação e selecção das unidades terminológicas constituem, para Wright (1997a:14), o primeiro passo na compreensão dos elementos específicos que constituem e enformam uma língua de especialidade. Este primeiro passo torna-se, depois, um elemento num processo contínuo de análise da evolução dessa língua e na compreensão da forma como verbaliza o conhecimento.

4. *Documenting terminological usage in domain discourse (within and across disciplines, within and across languages).*
5. *Support cross-cultural discourse (e.g. translation and parallel texts).* (Budin, 2006)

São estes aspectos que nos levam a considerar como necessária, enquanto parte do nosso método, a constituição e a existência de um *corpus*, preferencialmente do tipo comparável. Consideramos, no entanto, como Costa (2010) e Roche (2007, 2008), que o recurso ao *corpus* requer, para efeitos de análise e validação dos resultados obtidos, a presença de especialistas do domínio, num processo que, a nosso ver, deve assumir um cariz cooperativo e interdisciplinar.

Os critérios de selecção dos tipos de texto que vão compor um *corpus* devem, por isso, ser precisos e distintos conforme o tipo de investigação, a sua finalidade e âmbito de estudo, para que seja possível coligir e seleccionar apenas os dados essenciais com vista a formar a amostra necessária para aceder à informação sobre o conhecimento do domínio. O processo de compilação de um *corpus* exige, ainda, que se sigam certos princípios que apontam para a sua representatividade, exaustividade, homogeneidade e adequação, sempre associadas à pertinência e utilidade da tarefa a empreender e do domínio a representar.

Para que se atinjam essas metas com um alto grau de fiabilidade, a escolha dos textos deve ser feita em compatibilidade com os objectivos da pesquisa, para que se recolha e seleccione apenas o material necessário e se constitua a amostra desejável, sendo que um *corpus* textual, dependendo da função a que se destina, pode conter um ou mais tipos de texto, todos eles produzidos com propósitos e em contextos comunicativos diferentes, que importa ponderar⁹⁷. A recolha e constituição de um *corpus* de especialidade, representativo de um domínio, têm, assim, que ter em conta um conjunto de aspectos, de que se destacam o contexto profissional e cultural em que é produzido – contexto a que surge associado, de modo inextrincável, o significado.

⁹⁷ A estruturação de um *corpus* não está isenta de alguma opacidade, como se pode verificar pelas palavras de Rastier (2004): *Cependant, un corpus n'est pas plus un sac de mots qu'un nébuleux intertexte. Il est structuré d'une part en fonction d'une typologie des textes, qui se reflète dans leur codage, et d'autre part, dans chaque utilisation, par des sélections raisonnées de sous-corpus.*

Entendemos, tal como Leitner (2001: 151), que um *corpus* deve ser composto por vários tipos de texto e que a sua representatividade não é uma mera questão de dimensão. Um outro princípio importante é apontado por Biber (1993:256), para quem é essencial, no desenho de um *corpus*, que os parâmetros não sejam definidos à partida de modo estanque, a fim de que o *corpus* possa evoluir ciclicamente,

Assim, uma vez identificado o domínio específico e respectivos subdomínios, e delimitado o subdomínio a analisar, devem, na nossa perspectiva, ser definidas as características e critérios de selecção dos textos de especialidade⁹⁸ para constituir o *corpus*⁹⁹. No entanto, o conceito de texto de especialidade é, como descreve Costa (2005), demasiadamente genérico e complexo de forma a permitir um tratamento uniforme de todas as ocorrências de textos produzidos no seio de uma mesma comunidade científica¹⁰⁰. Os públicos a quem se dirigem, a pluralidade das situações e o enquadramento espaço-temporal em que os diversos textos¹⁰¹ são produzidos e consumidos são tão divergentes, que se impõe, para a autora, a constituição de uma tipologia textual, ponto de vista com o qual concordamos e que assumimos como necessário na aplicação do nosso método.

⁹⁸ Cabré (2007: 90) define os textos de especialidade pelas suas condições discursivas, cognitivas e linguísticas, condições que demonstram bem a natureza heterogênea dos textos de especialidade:

- *Condiciones discursivas: las propias del escenario especializado de este tipo de comunicación*
- *Condiciones cognitivas: el tema de qué tratan y la forma precisa de tratarlo*
- *Condiciones lingüísticas: las condiciones textuales generales (precisión, concisión y sistematicidad, las dos últimas en grados diversos según las condiciones discursivas), la forma textual macro y micro del texto, y sobre todo las unidades léxicas propias del dominio de que trata el texto.*

⁹⁹ Remígio (2010, 266) chama a atenção para a distinção entre texto e *corpus* que Sinclair (2004: 189) identifica ao afirmar que a distinção crucial não reside na “*amount of language it contains, nor is it the nature of the content, but the methodology, the way in which you approach it*”. Efectivamente, para Sinclair (2004: 189) a observação do conjunto de textos que integram um *corpus* não é directa, mas mediada por ferramentas específicas “*like query languages, concordancers, collocators, parsers and aligners*”.

¹⁰⁰ Para Kerremans (2005), a abordagem à classificação dos tipos de texto resulta normalmente numa taxonomia de tipos de texto. O problema deste tipo de classificação é, para o autor, o de que a noção de tipo de texto é “*of such a wide applicability that it can subsume a bewildering range of text-form variants (Emery 1991: 567). A second common type of classification – i.e. based on text function – is also problematic as several functions may be discerned in a text. Text classification may also be based on different communicative contexts (cf. Pearson 1998) or ‘situations’ such as province or domain (see e.g. Mason 1982)*”.

¹⁰¹ A propósito da distinção entre o conceito de texto e o de discurso, partilhamos da opinião de Costa (2005), para quem texto e discurso mantêm uma relação de interdependência, não podendo um ser pensado sem o outro uma vez que o discurso actualizado apresenta-se, geralmente, sob forma de texto. Falar em texto, implica a noção de discurso que foi proferido por um indivíduo, num espaço e num tempo específicos. Costa cita (Ricoeur, 1986: 137) que perspectiva o texto como «*[...] tout discours fixé par l’écriture*».

O estabelecimento de uma tipologia implica a classificação e a reunião de um conjunto de textos, que mantêm entre si relações “*de semelhança ao nível das respectivas macro e/ou microestruturas, sob uma mesma etiqueta, através da identificação de regularidades de um conjunto de textos, por oposição às regularidades de outros conjuntos*” (Costa, 2005)¹⁰².

Na nossa opinião, a existência de uma tipologia constitui um auxílio para o processo de selecção e compilação do *corpus*, quer no caso de um *corpus* especializado monolíngue quer no caso de *corpora* comparáveis, uma vez que permite agrupar e estruturar melhor o próprio *corpus* e atestar a sua representatividade e abrangência, sobretudo no caso de domínios complexos e interdisciplinares.

Contribui, também, para o processo de consulta, extracção e análise da informação terminológica, que pode, de acordo com as necessidades de pesquisa, incidir apenas sobre um dos tipos de textos, sendo os resultados obtidos mais contextualizados e mais facilmente acessíveis¹⁰³, auxiliando assim, por exemplo, na desambiguação de termos polissémicos ou na clarificação do uso desses termos, ao torná-los mais explícito.

No que se refere aos recursos que devem constituir o *corpus* somos de opinião que existe uma relação de complementaridade entre ambos. Os recursos linguísticos e semânticos constituem-se como elementos relevantes para a extracção e análise da informação do *corpus* textual, ao fornecerem pistas para o processo de extracção, elementos de comparação e, em última análise, de revisão e validação dos resultados obtidos. Por seu lado, os resultados obtidos a partir do *corpus* textual podem ajudar no processo de interpretação, delimitação e validação da informação disponibilizada por cada um dos restantes recursos, tornando-se, assim, em elementos relevantes para a construção e validação da representação conceptual.

¹⁰² Para a autora, é na “*identificação, designação e classificação das regularidades, que podem ser da ordem do sociológico, do psicológico, do histórico ou do linguístico, entre outros, que se situa a essência da problemática da tipologia*”.

¹⁰³ Esta situação pode acontecer, por exemplo, num domínio cuja actividade seja fortemente afectada por legislação específica, sendo assim necessário identificar e recolher textos do tipo jurídico, cujo conjunto poderá ser usado de modo independente para observar e validar candidatos a termos ou analisar o seu funcionamento e o seu significado em contextos específicos e situados.

3.4. Papel do terminólogo

No seio de uma rede colaborativa, a presença de um terminólogo pode funcionar como um elemento importante no apoio ao acesso, partilha e disseminação do conhecimento do domínio. Tendo em conta a situação sociocultural da rede, e os ambientes sociolinguísticos, macro e micro (Cabré, 1999: 18), em que o terminólogo actua, as suas funções e o grau de intervenção na análise e estruturação do conhecimento e na construção da ontologia, bem como na sua especificação multilingue, podem, no entanto, ter graus diferenciados.

Ao terminólogo cabe, tendo em conta as diferentes etapas do processo de conceptualização, preparar, facilitar ou conduzir o processo de eliciação e negociação do conhecimento entre os especialistas¹⁰⁴, por exemplo, através da identificação e disponibilização de recursos relevantes que apoiem esta negociação, promovam o consenso¹⁰⁵ - como recursos e fontes internacionais comumente aceites – e limitem as divergências semânticas entre línguas ou entre terminologias do domínio.

O especialista e o terminólogo estabelecem um diálogo e interagem num contexto definido, sendo, segundo Costa (2010: 05), a mediação entre o terminólogo e os especialistas feita através do texto oral, o que tem a vantagem de reunir os dois intervenientes numa situação sócio-pragmática precisa em que o contexto extralinguístico partilhado *“dissout les ambigüités et les reformulations discursives sont frequentes”*.

No entanto, este diálogo nem sempre é fluído. De facto, ao terminólogo não compete saber tudo¹⁰⁶. Tal como se deixa claro no ISO/TR 22134 (2007: 15), os terminólogos não podem prever todas as decisões. Podem errar nas suas escolhas terminológicas e

¹⁰⁴ O terminólogo é, por natureza das suas funções, um harmonizador (...) linguístico, por descrever o vocabulário técnico e científico, um harmonizador de conhecimento, por estruturar esse conhecimento e organizá-lo de modo a permitir a sua compreensão, e um harmonizador social, por interpretar a informação e facilitar a circulação da mesma em contextos comunicativos e sociais heterogéneos. (Romualdo, 2010: 06)

¹⁰⁵ Tal como refere o relatório técnico *Practical Guidelines for Socioterminology*, ISO/TR 22134 (2007: 15): *The speakers in a given professional medium will have a greater propensity to use the terms appearing in a standard if they know that the choice has been made by their peers. One of the possible responses to the question of implantability of the terminology is to prepare it with the users, as a function of their communication requirements.*

¹⁰⁶ Dubuc, em 1987, sintetizava assim o perfil do terminólogo: *“Fireman, safeguard, grammarian on duty, walking dictionary, language agent, gallicization force, high-quality communications promoter, a terminologist must be a bit of each...”*

reconhecer que os termos recomendados podem não vir a ser usados ou mesmo vir a ser rejeitados:

The procedures for drafting the terminology shall allow the choices to be reviewed and the terminology to be adjusted as a function of the reaction of the targeted users and, generally, as a function of the evolution of the usages.

Ao terminólogo compete, tendo em conta o contexto da rede e a situação comunicativa, entre as muitas tarefas possíveis, pesquisar e disponibilizar recursos terminológicos, lexicográficos, ontologias, mapas de conceitos, taxonomias, etc.; assistir no processo negocial; acompanhar as decisões dos especialistas; criar e actualizar o sistema de representação conceptual e auxiliar na sua implementação e, eventualmente, desempenhar também o papel de tradutor.

Deve, em particular, e de forma ágil, identificar os recursos apropriados à análise e uso de determinado conceito/termo, potenciando, assim, a partilha e a reutilização de conhecimento, tendo em especial atenção o modo como este processo se desenrola num ambiente multilingue e como contribui para a construção da ontologia em mais do que uma língua¹⁰⁷.

Romualdo (2010: 07) identifica, neste contexto, uma nova competência do terminólogo – a intercultural. Este, enquanto especialista na harmonização de informação e sua circulação, terá de ser um conhecedor das culturas com que contacta profissionalmente para poder desempenhar as suas funções de forma eficiente¹⁰⁸.

O trabalho de elaboração dos referenciais terminológicos, identificação de fontes e de análise do domínio é desenvolvido num ambiente, em parte, mais solitário, com uma participação menor do especialista nesta fase, cuja presença é constante no processo de

¹⁰⁷ Estas competências decorrem do carácter interdisciplinar da terminologia, que lida com aspectos da língua, da cognição e da comunicação em simultâneo.

¹⁰⁸ Como descrevem Lustig & Koester (2006: 02): “*Intercultural encounters are now ubiquitous (...) you no longer have a choice about whether to live or communicate among many cultures. Your only choice is whether you will learn to do it well*”.

conceptualização, nomeadamente durante a negociação da representação conceptual e do significado dos termos¹⁰⁹.

No que se refere ao uso da Web e à relação desta com as competências do terminólogo, esta relação comporta dois aspectos fundamentais: as competências que resultam da utilização da própria Web e as competências necessárias à utilização dos seus serviços no domínio da terminologia e da localização. Como descrevem Ciobanu & Varga (2010: 01), as ferramentas, os serviços e as novas formas de exploração da internet podem ser utilizadas na pesquisa terminológica, o que permitirá actualizar os métodos de trabalho neste domínio, promovendo a sua adaptação aos novos paradigmas informacionais.

Os autores identificam um conjunto extenso de novas competências, de que destacamos quatro, transcritas abaixo, pela relevância que possuem para o trabalho no contexto de redes colaborativas e para o papel do terminólogo enquanto promotor da partilha de conhecimento:

- Capacidade de assumir um papel activo nas comunidades virtuais, o que desenvolve a competência do terminólogo como intermediário eficiente no processo de mediação da comunicação especializada,
- Capacidade de colaboração no espaço virtual com grupos provenientes de diferentes culturas, o que obriga o terminólogo a desenvolver as suas competências de comunicação intercultural,
- Competências relativas à verificação e validação dos dados e informações provenientes da “inteligência colectiva”, devendo a escolha de fontes de informação ser feita com base num pensamento crítico,
- Competências comportamentais específicas ao contexto de trabalho e investigação, tais como a participação activa nas actividades de uma comunidade virtual, o desenvolvimento da confiança entre os membros de uma comunidade social ou a responsabilização do utilizador no que diz respeito às suas intervenções e à sua participação no espaço virtual (Ciobanu & Varga, 2010: 05).

¹⁰⁹ Na perspectiva de Costa (2010: 05), o trabalho com o especialista representa a situação mais desejável para um terminólogo, numa relação económica e produtiva em que, ao terminólogo, compete saber gerir o seu tempo e submeter ao especialista propostas de sistematização conceptual e denominativa.

Esta necessidade de novas competências e de novas abordagens à função do terminólogo vai para além do perfil clássico, acrescentando-lhe complexidade e novos desafios, em resultado da interacção contínua com o meio virtual e a comunidade de especialistas, o que leva a que pesquisa e o trabalho terminológico sejam, cada vez mais, actividades inter e multidisciplinares.

3.5. Função da língua inglesa na conceptualização e localização do domínio

A conceptualização de uma ontologia no seio de uma comunidade multilingue e multicultural acontece, não raramente, com recurso à língua inglesa como língua de trabalho, dada a predominância do seu uso nos contextos profissionais em termos globais. A existência de uma língua de trabalho corresponde, normalmente, aos requisitos e às necessidades próprias de redes colaborativas para o estabelecimento quer de objectivos e estratégias de desenvolvimento da rede quer de delimitação e análise do domínio – desempenhando a língua inglesa um papel central no processo comunicativo e na negociação e obtenção de consenso.

O recurso a esta língua acontece por diferentes motivos. O primeiro, directamente ligado ao funcionamento da rede colaborativa multilingue, e por uma questão de abrangência – o uso de uma língua bem aceite na comunidade e (re)conhecida pela maioria dos actores da rede para efeitos de comunicação contribui, até certo ponto, para ultrapassar as barreiras linguísticas e reforçar a rapidez e a qualidade da partilha de informação e conhecimento. O uso dominante da língua inglesa não significa, no entanto, que esta conduza, por si só, à criação de um entendimento comum, apesar de ser imprescindível para a sua concretização.

Outros motivos aparecem mais directamente ligados à capacidade de uso da informação e ao processo de localização. De facto, o uso de uma língua pivô introduz agilidade no processo de localização, i.e., localizar de uma única língua de partida (LP) para um número n de línguas de chegada (LC) é um processo reconhecidamente mais simples do que localizar de nLP para nLC .

Por outro lado, nem todas as línguas possuem ou oferecem o mesmo tipo de recursos (lexicais, terminológicos, entre outros), nem as ferramentas de tradução, automáticas ou de outro tipo, auxiliam no processo de localização com a mesma robustez e qualidade, pelo que partir da língua onde este tipo de recursos é mais abundante se torna importante para o desenvolvimento mais rápido e, até certo ponto, com um maior grau de fiabilidade¹¹⁰, porque mais apoiado, de ambos os processos.

Este carácter funcional que assume o uso da língua inglesa (ou de qualquer outra língua escolhida pela comunidade) no suporte à construção de uma representação conceptual no contexto de uma comunidade multilingue não lhe retira, como é natural, as características de ambiguidade que são inerentes a todas as línguas, nem elimina as dificuldades de a utilizar como veículo de comunicação e negociação durante aquela construção.

Tal como se especifica na figura seguinte, pretende-se que uma ontologia reflecta, em cada língua alvo, uma representação conceptual de alto nível, comum e partilhada. A representação conceptual inicial, conceptualizada e representada em língua inglesa¹¹¹, funciona, assim, como uma ontologia nuclear à qual se interligam as ontologias localizadas¹¹² nas diversas línguas naturais. Estas ontologias podem, se necessário, ser desenvolvidas posteriormente, com maior granularidade, de modo a integrar conhecimento mais específico de uma determinada cultura ou a responder a uma necessidade de partilha de conhecimento mais particular.

¹¹⁰ No ponto 4.2.1 descreveremos os recursos identificados como relevantes e que foram validados pelos especialistas, onde se verifica também a preponderância dos recursos em língua inglesa.

¹¹¹ O carácter de independência da estrutura resultante desta conceptualização e dos conceitos nela representados, pode, de acordo com Hubrich *et al.* (2008), não ser total, havendo, em alguns casos, a necessidade de recorrer a elementos adicionais que sustentem a representação dessa independência, como o uso de numeração para possibilitar a representação do conceito de modo independente da língua e da cultura em que é designado, ou o uso de normas internacionais como o SKOS, no caso da representação das relações conceptuais que suportem a construção de estruturas conceptuais independentes da língua.

¹¹² Esta metodologia de localização da ontologia permite, por um lado, uma actualização mais fácil das ontologias nas diferentes línguas e, por outro lado, num cenário de extracção de informação, facilita a navegação entre conteúdos de diferentes línguas.

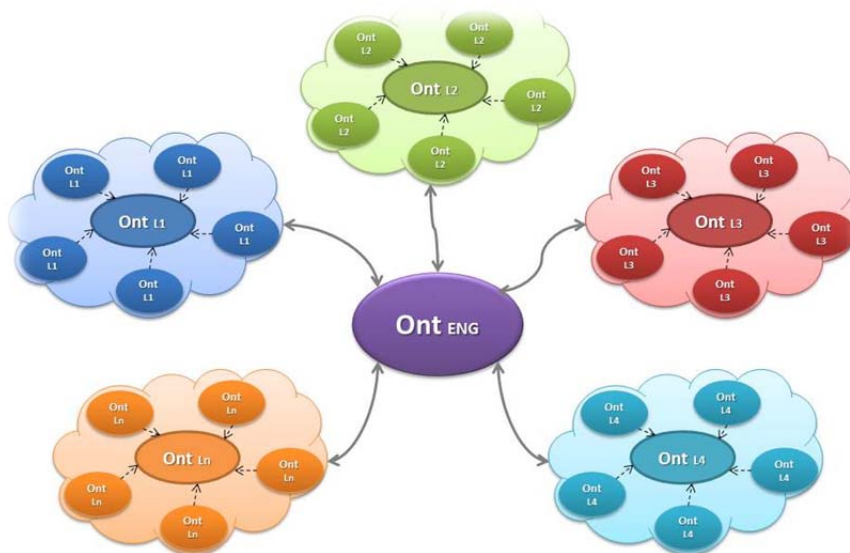


Fig. 10 – Interligação entre as ontologias em diferentes línguas

A ontologia de alto nível expressa em língua inglesa actua como elemento de interligação, pelo que a conceptualização inicial da ontologia deve ser cuidadosa, de modo a contribuir para a boa percepção do sistema conceptual do domínio.

A nossa proposta difere, assim, da abordagem proposta por Gödert (2008). O autor, tendo em conta o número crescente de sistemas que se procura integrar e tornar interoperáveis - mas que diferem linguística, estrutural e, até, tipologicamente, - desenvolveu um modelo destinado à construção do que denomina de um *“comprehensive international knowledge organisation system, involving a “core system” as the central element with the potential of interconnecting a multitude of various systems”*.

Este sistema central, que deve, necessariamente, ser de muito alto nível, funciona como um núcleo (*ontological spine*) e como uma língua de interligação (*switching language*) e oferece um eixo que interliga todos os sistemas. Para o autor, a ligação de cada sistema individual a este núcleo, como se se tratasse de satélites do sistema central, permite o acesso a cada um tendo como ponto de partida o núcleo.

No nosso método, a estrutura central – uma representação conceptual do domínio em língua inglesa - também pode desempenhar as funções de língua ou de núcleo de interligação, mas as restantes ontologias nas diferentes línguas representam conceptualizações que reflectem a inicial e não conceptualizações individuais acopladas.

Assumimos, assim, que a representação conceptual inicial, de alto nível, é passível de ser partilhada por diferentes línguas, podendo, no entanto, ser alvo de um processo de reconceptualização, em resultado do processo de localização e do *feedback* obtido a partir dos especialistas.

3.6. Contributo da representação do conhecimento

A comunicação é um acto deliberado dependente do uso de sinais utilizados como veículos na transmissão da mensagem pretendida; sinais que podem ser conjugados em sistemas de maior ou menor complexidade, em função do estágio de desenvolvimento da tecnologia e dos indivíduos envolvidos. No caso da comunicação especializada, provavelmente mais do que em qualquer outra situação, a comunicação supõe uma reflexão sobre a forma como se escolhe a linguagem, como se constrói a mensagem e como se a transmite.

A estruturação da representação conceptual é um factor essencial na apropriação e transmissão do conhecimento em qualquer domínio de especialidade, o que pressupõe uma avaliação por parte do especialista no processo de identificação e a delimitação dos conceitos, tarefa que não é isenta de dificuldades. Com efeito, é expectável que os especialistas de um domínio possam ter vários pontos de vista no que respeita:

1. à percepção do domínio, à sua natureza, estrutura e ao modo como representar o conhecimento do domínio,
2. à escolha dos conceitos relevantes para representar e delimitar o domínio,
3. à terminologia a usar para o descrever e ao significado dos termos,
4. à definição das relações relevantes entre os conceitos e à sua validade lógica,
5. ao modo como são designados noutras línguas.

Dada a diversidade de pontos de vista, os termos devem ser alvo de um processo de organização em sistemas estruturados que reflectam uma organização conceptual, correspondendo a delimitação do domínio a uma função das necessidades de estruturação do conhecimento e das necessidades expressivas dos vários utilizadores envolvidos.

Cada área de especialidade pode composta por um conjunto de domínios, sendo uma das tarefas quer da gestão terminológica e quer da gestão do conhecimento a representação conceptual dos diferentes (sub)domínios, a qual deverá reflectir a visão, preferencialmente partilhada, dos especialistas sobre o seu domínio de especialidade.

A fase inicial de recolha e análise da informação e criação de uma representação do conhecimento é complexa, podendo esta complexidade aumentar quando estamos perante equipas multinacionais, em resultado das características heterogéneas dos actores, das suas experiências profissionais e culturais, entre outros factores. No processo de conceptualização de um domínio - processo cíclico e iterativo, resultante do *feedback* obtido da negociação e interacção -, uma disposição e uma visualização estruturada e consensualizada dos conceitos, obtida com recurso ao uso de mapas conceptuais, pode, neste contexto, apresentar vantagens na melhoria da eficiência dos processos de análise, estruturação e representação do conhecimento do domínio, bem como do(s) sistema(s) em que o resultado final será integrado e utilizado.

Para Pereira (2010: 110), a utilização de mapas de conceitos permite a utilizadores inexperientes, que não estão preparados para lidar com as restrições da semântica formal, concentrarem-se na tarefa de construção do modelo conceptual num ambiente sem restrições. O objectivo é permitir aos utilizadores (todos os participantes no processo de conceptualização) iniciarem a construção do modelo informalmente, sem terem de traduzir o seu *know-how* para qualquer formalismo de representação de conhecimento. A experiência mostra, segundo a autora, que os mapas de conceitos são uma técnica adequada para ser usada durante o processo de conceptualização colaborativa¹¹³.

De acordo com Kelly (2009: 12), sempre que duas (ou mais) culturas diferentes se encontram, para que exista entendimento, é necessário criar-se um terceiro espaço: *“in order for meaning to exist it is not sufficient two opposite parties meeting; there has to be*

¹¹³ O processo de conceptualização recorre, para além da representação em mapas conceptuais, a diferentes modelos. No entanto nenhum sistema de representação está livre de factores problemáticos, como salienta Wright (2007: 175), os quais criam obstáculos à sua universalidade e interoperabilidade, por um lado, porque o tipo de estratégia de representação adoptada pode levar a que as estruturas conceptuais não sejam reconhecidas por outros sistemas e, por outro, porque os preconceitos linguísticos e culturais dos criadores dos sistemas de organização do conhecimento (KOS) e dos sistemas de representação do conhecimento (KRR) tendem a reflectir-se no sistema. Para a autora, até mesmo grupos de trabalho multiculturais e conscientemente orientados para o consenso produzem representações que apenas reflectem a sua visão colaborativa do mundo, que pode não ser transparente para utilizadores fora desse grupo.

some third space, in which they have something in common to be able to meet". Este terceiro espaço ganha, inicialmente, forma no nível mais detalhado de interação, alargando-se, de seguida, aos estratos sociais e às lógicas funcionais dos grupos que os constituem, podendo ser aplicado à lógica interna da comunidade e das relações empresariais.

Tendo por base esta ideia, e de modo a criar este "terceiro espaço", a abordagem à sistematização e representação do conhecimento do domínio será conduzida, no contexto do método, através da sua representação em mapas conceptuais¹¹⁴, que assumirão, um carácter de sistemas de natureza semiformal de suporte ao processo de conceptualização e localização.

Os mapas conceptuais constituem, na nossa opinião, uma ferramenta intuitiva que pode ser usada com sucesso para explicitar e comunicar o conhecimento do domínio, permitindo a sua captura e representação de modo acessível aos diferentes actores de uma rede ou comunidade¹¹⁵. De acordo com Tricot e Roche (2006: 1), o uso do que designam de *cartografia semântica*¹¹⁶, tem como objectivo mapear o conhecimento de um domínio através do uso da sua estrutura semântica para partilhar e criar novo conhecimento. Para os autores, a visualização do conhecimento é vista como o uso de representações visuais de dados abstractos *"to amplify cognition to creative ends and/or knowledge sharing"* (Tricot & Roche, 2006: 02).

¹¹⁴ Um mapa conceptual é um recurso de representação esquemática de significados conceptuais, organizado através de uma estrutura bidimensional de proposições, sendo que entendemos proposições como: *"statements about some object or event in the universe, either naturally occurring or constructed. Propositions contain two or more concepts connected with other words to form a meaningful statement"* (Novak, 2003: 1). A(s) palavra(s) de ligação especifica(m) a relação entre os dois conceitos, atribuindo um significado à relação. Os mapas constituem uma ferramenta usada para sintetizar e consolidar a informação a partir de diferentes fontes de pesquisa e para 'simplificar' a abordagem a problemas complexos.

¹¹⁵ Como refere Costa (2009) a comunicação especializada diz respeito à transmissão de conhecimentos, *"havendo um intercâmbio de formas e de representações entre o emissor e o receptor. Embora seja o emissor a transmitir o conhecimento, impregnado da sua visão do mundo, é o receptor que filtra a mensagem e a reconstrói em função do seu conhecimento"*, pelo que importa utilizar um meio de representação do conhecimento acessível a ambos que evite a incompreensão do sentido da representação. Estabelecemos aqui um paralelismo com Cabré (1999), que adverte para o facto de que quando os utilizadores (de uma terminologia) não estão bem definidos ou quando o propósito da terminologia é da expansão do uso das *"proposed forms (...), the way the term collections are presented is almost as important as the content of the material itself"*.

¹¹⁶ *"Semantic cartography is a modelling-oriented visualisation of an informational space, i.e. a visualisation guided by the semantic structure of the information space"*. (Tricot & Roche, 2006: 01)

Os mapas conceptuais, entendidos como uma representação possível de uma dada estrutura conceptual, possuem, na nossa perspectiva, como característica essencial, um carácter dialógico, i.e., servem de base a uma negociação e a um diálogo sobre o valor da representação conceptual, cuja existência deriva da própria estrutura do objecto de conhecimento em análise. São estruturas dialógicas conceptuais de carácter semiformal, e permitem uma análise das diferentes propostas de modelação do domínio e da sua evolução, apoiando o processo de percepção e representação iterativo e partilhado do domínio e facilitando o trabalho de formalização subsequente.

3.7. Ferramentas e ambientes de suporte ao desenvolvimento do método

A implementação do método implica o desenho e a definição de um ambiente de apoio ao seu desenvolvimento e a selecção de um conjunto de ferramentas que tornem as actividades de localização e eventual reconceptualização viáveis.

Assim, no que respeita à definição do ambiente de trabalho, decidimo-nos pelo uso de mapas conceptuais, como descreveremos no ponto seguinte. Estes podem, na nossa perspectiva, constituir-se também como uma ferramenta que permite que o processamento multilingue da informação, com recurso à localização, aconteça no momento da conceptualização, ao possibilitarem aos especialistas a inserção dos equivalentes dos conceitos e das relações conceptuais, por um lado, e por outro, uma correcta avaliação e validação desses mesmos equivalentes, uma vez que surgem situados na estrutura conceptual.

A implementação de um método de apoio à localização implica, por outro lado, a selecção prévia de um conjunto de aplicações para auxiliar na pesquisa e selecção dos equivalentes. A escolha de uma determinada aplicação para acompanhar um processo de localização depende, como veremos no ponto 3.6.2, da consideração de um conjunto de factores e requisitos, nomeadamente, o tipo de conteúdo a ser localizado, o contexto em que o processo decorrerá, os instrumentos e recursos ao dispor dos elementos mais directamente envolvidos no processo e os objectivos a atingir com o processo nas suas diferentes etapas.

3.7.1. Espaço de conceptualização e localização

A representação visual não-verbal a partir de mapas conceptuais não é dependente de uma única língua, o que não exclui o facto de ser regida por regras e convenções socioculturais. É, no entanto, desenvolvida a partir da acção conjunta de um dado número de especialistas, pelo que reflecte, necessariamente, a sua visão do mundo.

De modo a tornar o processamento multilingue da informação possível a partir de mapas conceptuais e a apoiar os diversos actores nesse processo, procurámos criar um ambiente de trabalho baseado nas funcionalidades oferecidas pela ferramenta *CmapTools* que promovesse a localização na fase de conceptualização. Procurámos que esse interface fosse de uso simples e adequado ao tipo de utilizadores.

Para tal, desenvolvemos um cenário de trabalho colaborativo que viabilizasse a localização das representações conceptuais num ambiente multilingue e definimos o conjunto dos elementos linguísticos e terminológicos que consideramos suficiente para apoiar a localização neste ambiente. Assim, cada representação conceptual será composta por um conjunto de elementos de base terminológica, nomeadamente:

1. Cada conceito será definido em língua natural. As definições usadas devem ser dependentes do domínio e obtidas a partir de recursos validados e aceites pelos especialistas (ex.: normas, classificações). A definição será expressa em língua inglesa como comentário e visualizável para o utilizador no momento da conceptualização, tal como se demonstra na figura seguinte:

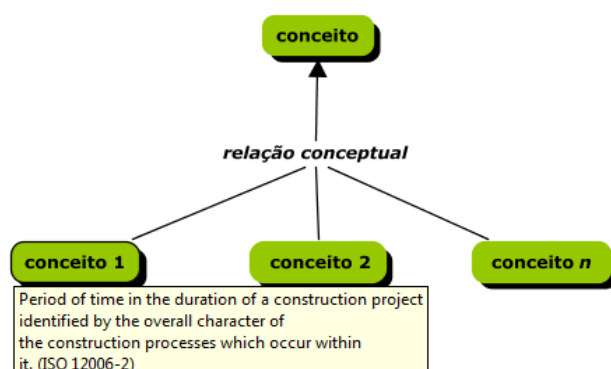


Fig. 11 – Espaço de conceptualização e localização – acesso a definições

Pressupondo que o domínio está identificado e delimitado, acreditamos que a existência de definições em língua natural, imediatamente acessíveis ao especialista, e com a sua origem identificada, apoiam-no na conceptualização e localização, uma vez que funcionam como um suporte à aquisição de conhecimento e à negociação da representação conceptual ao permitir uma mais efectiva:

1. distinção entre conceitos,
2. desambiguação,
3. noção do significado do conceito,
4. pesquisa e selecção do equivalente em falta,
5. avaliação/validação do equivalente proposto.

O papel que as definições¹¹⁷ desempenham é importante ao longo de todo o processo de conceptualização, uma vez que, se elaboradas coerentemente¹¹⁸, agem, por um lado, como um indicador do tipo de relações que se estabelecem entre os conceitos do domínio, quer das relações hierárquicas quer das não hierárquicas e, por outro, permitem determinar se o conceito descrito corresponde ao objecto.

Este ambiente permite a inclusão de mais do que uma definição, que podem ser de diferentes tipos - ex.: definições operacionais sobre fórmulas, definições processuais - de modo a não limitar o especialista no momento de seleccionar a informação a fornecer e a permitir que este contribua efectivamente para a conceptualização, independentemente do seu tipo e nível de especialização.

¹¹⁷ Uma definição permite, segundo Sager (1990: 21), fixar a referência precisa de um termo a um conceito e ao mesmo tempo criar e, por isso, declarar relações com outros conceitos dentro de uma estrutura de conhecimento. Para o autor (1990: 51), a definição: *“is not alone in providing a link between term and concept but is supported by the declaration of a sufficient and appropriate set of conceptual relationships. The definition can therefore concentrate on the essential characteristics a concept has in common with others and which differentiate it from other concepts; by contrast the relationships indicate the type of link a concept has to other concepts in the system”*.

¹¹⁸ Uma definição deve, segundo Cabré (1999: 106), descrever o conceito e reunir as dimensões relevantes para cada domínio de especialidade e *“be located in the perspective of the conceptual field a concept belongs to; be appropriated for the aims of the project in which they are presented”*.

2. Através do uso do sistema de comentários, permitir a inserção e o acesso dos diferentes especialistas aos equivalentes nas diferentes línguas de trabalho, tal como se representa na figura seguinte:

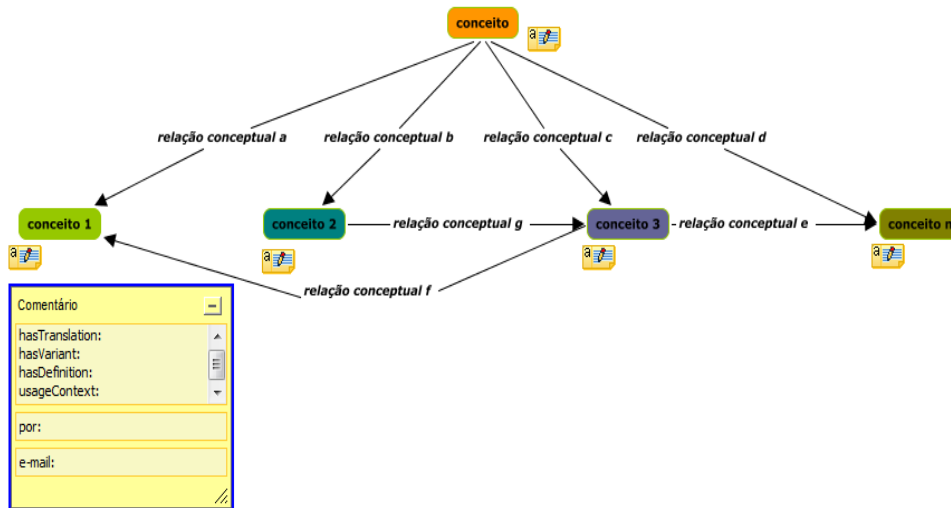


Fig. 12 – Espaço de conceptualização e localização – escolha de equivalentes

Para tal foi criado, recorrendo ao sistema de comentários do *CmapTools*, um ambiente com o seguinte conjunto de elementos:

<p>1. Definição de equivalentes, com a possibilidade de inclusão de:</p> <ul style="list-style-type: none"> a. <i>equivalente</i> b. <i>uma ou mais variantes</i> c. <i>definições</i> d. <i>contexto de uso</i> 	<p>1. Elementos disponíveis para preenchimento: (<i>LMF FRAMEWORK</i>¹¹⁹)</p> <ul style="list-style-type: none"> 1. <i>hasTranslation:</i> 2. <i>hasVariant:</i> 3. <i>hasDefinition:</i> 4. <i>usageContent:</i>
--	--

Estes elementos pretendem permitir o acesso aos elementos terminológicos e linguísticos do domínio. A sua inclusão teve por base as categorias propostas na norma ISO

¹¹⁹ <http://www.lexicalmarkupframework.org/>

*Lexical Markup Framework (LMF)*¹²⁰. São, no contexto de uma rede colaborativa e para os propósitos de conceptualização e localização, os elementos que consideramos suficientes para o desenvolvimento do processo. Pretendemos, com a inclusão destes elementos, permitir a gestão da informação multilingue no contexto de um mapa conceptual e criar a possibilidade de acesso à informação existente nas diferentes línguas de trabalho ao conjunto dos participantes neste processo.

O recurso a um número reduzido de elementos tem em conta os limites temporais a que os processos de conceptualização e localização estão normalmente sujeitos e procura criar um ambiente simples e funcional que promova a participação do especialista. Este pode inserir equivalentes, variantes, definições, e contextos de uso na sua língua de trabalho (ou na que considerar necessário/conveniente) na zona de comentários, de modo a sustentar a sua escolha/proposta. Admitimos, assim, neste espaço, a existência da sinonímia, por uma questão de funcionalidade e de flexibilidade do processo e como modo de promover a compreensão e a comunicação.

Este processo pode decorrer de dois modos: 1) a partir da sugestão prévia de equivalentes elaborada pelo terminólogo, obtidos com base nos recursos de especialidade multilingues previamente identificados e validados (dicionários, terminologias, classificações, tesouros ou mesmo ontologias), sendo estes equivalentes revistos e validados pelos especialistas, ou 2) a partir da sugestão feita pelo próprio especialista, com base na sua experiência e no conhecimento que possui da sua língua de especialidade.

O especialista deverá, sempre que possível, identificar os recursos que o auxiliaram no processo; recursos que podem ser adicionados ao *corpus* para consulta futura. Na nossa perspectiva, sempre que estivermos perante um caso de reutilização de recursos de conhecimento cujo formato inicial é multilingue, nomeadamente ontologias, tesouros ou classificações anteriormente traduzidas, o equivalente proposto pode ser o da fonte original, decorrido um processo de validação inicial, - desenvolvido com a participação do especialista -, da sua adequação ao domínio e aos propósitos da conceptualização.

Finalmente, e no que respeita às relações conceptuais, estas podem também ser localizadas, surgindo o equivalente sob a forma de informação no nó da relação. Na nossa

¹²⁰ <http://www.isocat.org/interface/index.html>

perspectiva, e para os efeitos do processo de localização da representação conceptual, devem ser mantidas as relações conceptuais utilizadas na representação conceptual inicial, uma vez que foram já alvo de análise e validação. Contudo, estas devem ser sempre ser alvo de um processo de verificação no que respeita à sua validade lógica e à adequação ao processo de pesquisa e extracção de informação.

Concordamos, no entanto, com Hubrich *et al.* (2009) quando afirmam que a localização, a longo prazo, envolve também o desenvolvimento de um conjunto de relações tipo (*typed relations*) a partir da tipologia de relações do sistema de gestão do conhecimento particular de modo a criar uma ferramenta eficiente para a navegação e exploração do conhecimento. Para os autores, e numa situação ideal, a maior especificidade inerente ao sistema de conhecimento localizado deveria reflectir-se também na especificação de um conjunto de relações apropriadas à representação local¹²¹.

A localização dos termos que designam os conceitos e das relações que os ligam para uma língua específica representa uma condição essencial ao seu uso nos sistemas de gestão do conhecimento, quer num contexto internacional quer num contexto local. A localização implica que se integrem os conceitos da representação conceptual inicial na representação conceptual local, sem a perda ou alteração do seu significado, pelo que, na nossa perspectiva, a possibilidade da presença “activa” das diferentes línguas reforça o processo de construção da própria conceptualização e a sua granularidade ao permitir um entendimento comum dos conceitos nela representados e do seu significado.

O recurso a este tipo de visualização permite que se implementem os princípios da univocidade, da facilidade de compreensão, da transparência e da possibilidade de ampliação, apontados por Arntz e Picht (1995: 105-106) como essenciais para a construção de um sistema conceptual, dado que estes mapas reflectem, até certo ponto, princípios e estruturas aceites entre a comunidade de cientistas ou técnicos, pela sua grande

¹²¹ Na nossa opinião, a existência de ontologias que pretendam especificar dois níveis diferentes de conhecimento - um global e um local – coloca um desafio especial para a construção das estruturas semânticas locais, tornando-se essencial perceber as diferenças (distâncias) que poderão existir entre a estrutura global e a local de modo a desenvolver um conjunto de relações conceptuais que permita o estabelecimento da necessária interoperabilidade semântica entre as ontologia de diferente nível. O desenvolvimento do processo de conceptualização de uma ontologia está, assim, ligado à definição de um conjunto de relações conceptuais que sejam suficientemente abrangentes e relevantes para o domínio e que permitam a sua reutilização.

universalidade e pelo seu carácter não-linguístico, o que facilita a sua compreensão e cria um contexto comunicativo que agiliza a comunicação.

A interpretação da representação e da informação contida nos mapas conceptuais depende, no entanto, de factores como o público-alvo; o propósito da representação; o nível de conhecimento e os perfis dos diferentes interlocutores e a familiaridade com as convenções do domínio. Os diferentes interlocutores devem, preferencialmente, possuir um sistema de referências comum, para que a comunicação exista e para que a negociação se torne viável.

Aos especialistas que desenvolvem a representação compete, então, conhecer, por um lado, as regras gerais de percepção visual e, por outro lado, as experiências visuais comuns, em termos de estruturação do conhecimento, que os seus interlocutores possuirão, de forma a poderem comunicar com eficiência e validar as propostas dos seus parceiros. A construção desses modelos de modo consistente representa, assim, um desafio, sobretudo quando se pretende seguir um formalismo de representação semiformal de estruturas conceptuais.

3.7.2. Aplicações e recursos linguísticos de suporte à localização

A definição de uma estratégia de localização depende do modo como percebemos, por um lado, o conteúdo a localizar – no nosso caso, uma ontologia – e, por outro, os diferentes interlocutores que participarão no processo; o contexto em que se inserem e que sustentará o desenvolvimento do processo; os limites temporais a que este está sujeito; e o processo definido para a obtenção de *feedback* e para a sua integração.

Para que a localização decorra num ambiente que permita aos elementos de uma rede uma participação efectiva e eficiente, torna-se, assim, necessário definir um conjunto apropriado de ferramentas que possa lidar com uma tipologia mista de recursos monolíngues, bilingues e multilingues.

A esta definição surgem associadas um conjunto de dificuldades, uma vez que os actuais serviços de tradução são quase exclusivamente focalizados na tradução de documentos, não prevendo a participação dos utilizadores finais, nem disponibilizando

aplicações que auxiliem a ultrapassar o problema das barreiras causadas pela presença de diferentes línguas de trabalho num mesmo espaço colaborativo e de partilha do conhecimento. Estas ferramentas não incorporam, tradicionalmente, as necessidades de comunidades que funcionam em rede, nem prevêem a participação do (potencial) utilizador final no processo.

Estas condições implicam a necessidade de disponibilizar um espaço de trabalho que promova a localização da representação conceptual num contexto que permita a comunicação entre os interlocutores; espaço que considere a estrutura social e organizacional da comunidade, bem como o tipo de competências existentes.

Para tal, será necessário identificar e considerar os ambientes e fluxos de trabalho já desenvolvidos e definir um que disponibilize um conjunto de ferramentas e recursos linguísticos mínimos e facilmente acessíveis, que promova a participação dos diferentes interlocutores e línguas e que funcione como uma solução de auxílio ao processo de localização orientada para os requisitos idiossincráticos da rede.

Sendo a localização uma actividade baseada no conhecimento (Wilss, 1996), a selecção de técnicas, métodos e ferramentas apropriadas à localização dos termos da ontologia, depende, no entanto, dos recursos disponíveis para cada língua em particular e para o domínio a ser representado. Como é previsível, quer o número quer o tipo de recursos (ex.: terminológicos, lexicográficos) diferem bastante dependendo da língua de origem ou de destino, destacando-se, como referimos já, a oferta de recursos em língua inglesa.

O mesmo acontece com as ferramentas de tradução e de localização que se encontram em estádios de desenvolvimento diferentes e cujos resultados continuam a ter graus díspares de granularidade, de qualidade - em muitos casos insatisfatória e causadora de ruído - e de interesse, dependendo do par de línguas ou da aplicação em uso, pelo que o seu uso deve ser analisado criticamente à luz das necessidades de localização de uma língua de especialidade.

Propomos, assim, que o processo de localização da ontologia, executado a partir da representação conceptual em mapas de conceitos, seja assistido pelo recurso a um conjunto de ferramentas cujas características auxiliem e promovam o acesso e o trabalho num

ambiente multilingue de conceptualização, ferramentas como sistemas em linha de tradução automática (ex.: *Google Translator*¹²², *Microsoft Bing Translator*¹²³, ou o *Linguee*¹²⁴), bases de dados terminológicas (ex.: *IATE*), bases de dados lexicais (ex.: *Wordnet*¹²⁵), dicionários (ex.: *Wiktionary*¹²⁶, *OmegaWiki*¹²⁷) e outros recursos mais genéricos como o *LazyTerm*¹²⁸, que poderão, no entanto, não contemplar ainda algumas das línguas para as quais a ontologia deverá ser localizada.

Foram considerados outras ferramentas¹²⁹ como o *EuroWordNet*¹³⁰ ou *Babelnet*¹³¹, mas a utilização implicaria instalações locais, sendo o seu uso de maior complexidade para os utilizadores. Por outro lado, não abrangem uma parte das línguas que nos interessam mais directamente, nalguns casos também a portuguesa, sendo os resultados que obtivemos no seu uso comparativamente menos interessantes e completos do que os das aplicações acima identificadas.

Uma das razões para a escolha e uso daquelas aplicações prende-se com o grau de facilidade de acesso e o uso intuitivo que as ferramentas oferecem aos diferentes utilizadores sem competências específicas de tradução ou localização. Como indica Kargioti (2010: 02), o carácter intuitivo e a usabilidade das aplicações do tipo *Web 2.0*, *“further enhanced with semantic technologies for “intelligent” information processing, and combined with automated translation services, can open participation to individuals regardless of educational, professional or language background”*.

Naturalmente que o recurso a este tipo de ferramentas tem desvantagens, nomeadamente as oriundas do facto de alguns dos seus utilizadores não possuírem

¹²² <http://translate.google.pt/>

¹²³ <http://www.microsofttranslator.com/>

¹²⁴ <http://www.linguee.com/>

¹²⁵ <http://wordnet.princeton.edu/>

¹²⁶ <http://www.wiktionary.org/>

¹²⁷ http://www.omegawiki.org/Meta:Main_Page

¹²⁸ <http://terminotrad.info/RogerMcKeon/LazyTerm/TerminoParesse-Rev.3e.html>

¹²⁹ Uma dessas possibilidades passa pelo recurso e consulta a sítios multilingues de referência para o domínio que possuam informação facilitadora da identificação e selecção dos equivalentes.

¹³⁰ <http://www.illc.uva.nl/EuroWordNet/>

¹³¹ <http://lcl.uniroma1.it/babelnet/>

conhecimentos específicos quanto ao seu uso ou não possuírem formação ou competências mínimas em tradução, como já dissemos, pelo que se torna necessário desenvolver e fornecer um conjunto de orientações que promovam o seu uso efectivo e o processo de pesquisa e selecção da informação. Estas orientações devem considerar os propósitos da rede colaborativa; o perfil dos seus utilizadores; as suas necessidades de acesso à informação multilingue; o contributo expectável dos utilizadores e, finalmente, o domínio a ser representado e localizado.

A escolha de um conjunto de ferramentas e recursos que tenha em consideração os elementos descritos implica que se explore uma abordagem conducente à obtenção de uma localização com qualidade, que assentará necessariamente na validação dos resultados obtidos com base na colaboração dos especialistas de domínio, cuja intervenção e papel serão descritos com mais pormenor no ponto 3.7.

Na aplicação do método não consideramos, para efeitos de localização, todos os elementos linguísticos que constituem a ontologia (ex.: as definições), uma vez que o processo de localização centrar-se-á ao nível dos termos que designam os conceitos da ontologia.

Os métodos de localização, no caso específico de uma rede colaborativa têm, assim, que se adaptar aos actores dessa mesma rede e ao modo como se espera que contribuam para o processo de conceptualização e localização. Para tal, existem diferentes possibilidades, como o recurso à tradução directa, caso se verifique a existência de recursos fiáveis, nomeadamente glossários e dicionários bi ou multilingues de especialidade ou tesouros específicos do domínio que abranjam as línguas de trabalho da rede. Na falta deste tipo de recursos, poder-se-á considerar o recurso a outras técnicas, como as baseadas na tradução automática ou em *corpora*.

Do mesmo modo, deve-se ter em conta os diferentes momentos já descritos, e as diferentes necessidades de localização a que poderão conduzir, uma vez que, numa primeira fase, a língua de partida é o inglês, língua utilizada na rede para efeitos de comunicação, negociação e estruturação do conhecimento e, numa segunda fase, a língua de partida poderá vir a ser uma das línguas de trabalho dos utilizadores da rede.

3.8. Fluxo de trabalho do método

O desenho do método teve em conta, como descrevemos, o contexto de estruturação e de uso do conhecimento, os objectivos a atingir com esse uso, e os meios disponíveis para a apreensão da complexidade e riqueza semântica do domínio e para a sua representação em termos multilingues (ponto 3.2) numa ontologia de domínio.

O método para a especificação multilingue de ontologias foi concebido de modo a ser integrador e criar um fluxo de trabalho que conduzisse a um enriquecimento linguístico gradual das representações conceptuais e que permitisse e promovesse a iteração constante entre especialistas e recursos lexicais, terminológicos, entre outros; a iteração com aplicações de tradução e localização; bem como a iteração entre os especialistas e entre estes e o terminólogo.

Tal como se sistematiza na figura seguinte, o método propõe diferentes fases conducentes à construção de uma ontologia representada em mais do que uma língua.

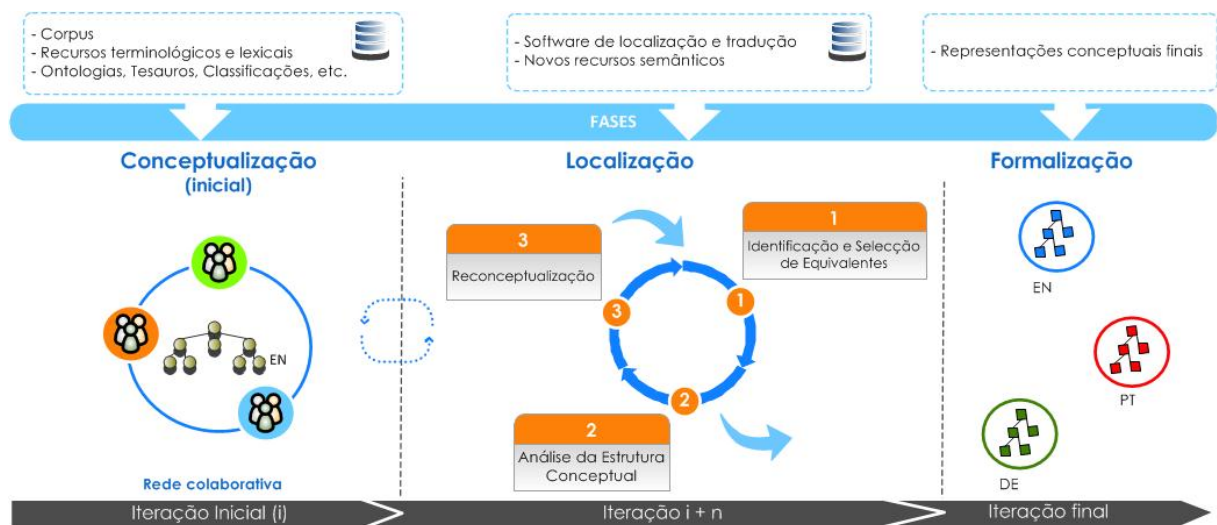


Fig. 13 – Fluxo de trabalho do método para a especificação multilingue de ontologias

Na primeira fase da construção da ontologia procede-se à construção de uma representação conceptual inicial do domínio, construída numa língua de trabalho previamente acordada (normalmente a língua inglesa), a partir da qual se desenvolve o enriquecimento multilingue da representação conceptual. Partimos, assim, de uma

representação partilhada e consensual do conhecimento do domínio, expressa numa língua de partida, para determinar o modo como é expresso nas várias línguas de chegada. Esta conceptualização tem como *input* (1) os contributos dos especialistas do domínio (ponto 3.9) e (2) os resultados da recolha e análise de um *corpus* específico do domínio (ponto 3.3.1), composto por recursos representativos das diferentes línguas, destacando-se, neste processo, o papel da reutilização e adaptação de recursos linguísticos e semânticos seleccionados como relevantes para o domínio.

Construída a representação conceptual inicial, inicia-se uma segunda fase. Nesta fase desenrolam-se os processos de localização e de (re)conceptualização, processos que devem ocorrer consecutivamente ao longo de toda esta fase. Como afirmámos, o carácter iterativo e, até certo ponto, cíclico, proposto para o desenrolar dos processos de localização e (re) conceptualização pretende possibilitar uma partilha de conhecimento mais rica e uma análise mais aprofundada da informação disponível, bem como a eventual reconceptualização da representação conceptual inicial do domínio.

Nesta fase é proposto um fluxo de trabalho e são identificadas e disponibilizadas um conjunto de ferramentas de auxílio à localização, de modo a assistir os especialistas e restantes interlocutores no desenvolvimento do processo e na identificação e selecção dos equivalentes. Para tal, como descrevemos antes (ponto 3.7), é disponibilizado aos especialistas um espaço virtual de conceptualização e de localização, de modo a permitir e sustentar a situação comunicativa específica.

Este espaço disponibiliza ferramentas de trabalho para apoiar o acesso ao conhecimento e à sua representação em termos multilingues, através do recurso a mapas conceptuais, que podem ser acedidos de forma partilhada, nos quais os especialistas representam os conceitos do domínio e as relações que estabelecem entre si, bem como os seus equivalentes nas línguas de chegada. O espaço dá ainda a possibilidade aos especialistas de acrescentarem outra informação relevante para a análise da conceptualização do domínio e para a sua localização, como definições em língua natural, variantes e contextos de uso.

Nesta fase também, o especialista pode seleccionar e propor novos recursos semânticos que considere relevantes para a análise e validação do sistema conceptual, podendo estes podendo funcionar como *input* para o desenvolvimento de representações

de conhecimento mais específicas ou como meio para a identificação e elicitación de conceitos ainda não representados. Estas propostas são alvo de uma análise e validação por parte dos restantes especialistas.

Após o processo de validação das propostas de reconceptualização e de localização, que poderá envolver o conjunto dos actores afectos ao processo ou apenas uma parte destes, em função das necessidades e do fluxo de trabalho pré-determinado para a rede colaborativa ou para a própria construção da ontologia, tem início a terceira e última fase.

Nesta procede-se à especificação formal das representações conceptuais do domínio e, finalmente, à sua implementação e ao seu teste, tendo em conta os objectivos de acesso e gestão do conhecimento previamente definidos. Esta última fase poderá, eventualmente, conduzir a uma nova iteração, em função dos resultados obtidos durante o teste e/ou o uso da ontologia.

3.9. Papel do especialista de domínio

Os especialistas de uma rede colaborativa são, cada vez mais, convidados a desempenharem um papel activo no processo de conceptualização do domínio. Dadas, neste caso, as características e os propósitos multilingues desse processo, acreditamos que esse papel será, em parte, extensivo ao processo de localização.

A especificação multilingue de uma ontologia diz respeito à localização de um artefacto que é, em si mesmo, uma representação estruturada do conhecimento específico de um domínio, pelo que estamos perante um objecto que deve ser entendido como um todo, tendo, por isso, o processo de localização que se centrar não só nos elementos individuais que constituem o sistema conceptual e na procura dos seus equivalentes, mas antes, e sobretudo, na estrutura conceptual como um todo e aferir se esta representa o conhecimento tal qual é percebido pela comunidade para a qual está a ser localizado.

Esta nova centralidade – que vai para além da simples procura de equivalentes, e destaca o sistema conceptual ainda em construção como elemento determinante – coloca os especialistas das diferentes comunidades linguísticas que participam na rede colaborativa num plano de relevo, uma vez que funcionarão como agentes com uma dupla função: por

um lado, a de validar os equivalentes seleccionados e auxiliar no processo de desambiguação; por outro, a de verificar se o sistema conceptual daí resultante representa, efectivamente, o conhecimento de domínio tal qual é percebido, na sua língua de trabalho, pela sua comunidade de especialistas.

Estes actores tornam-se, assim, em paralelo, produtores e utilizadores de um artefacto semântico multilingue¹³², tal como se representa na figura seguinte; artefacto que representa o conhecimento do domínio e que é desenvolvido de acordo com um conjunto de critérios e propósitos bem estabelecidos, no contexto de um projecto e de uma rede e de acordo com as limitações sócio-funcionais e temporais inerentes à sua conceptualização e especificação formal.

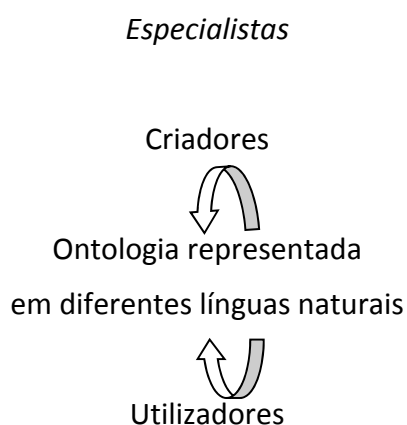


Fig. 14 – Papel dos especialistas

Os especialistas possuem, no nosso ponto de vista, os dois tipos de conhecimento identificados por Rothkegel (2003) como sendo relevantes também para a tradução e a localização: conhecimento do domínio (*world knowledge*) que se refere às estruturas conceptuais dos objectos e estado das coisas, e conhecimento do texto, referente aos padrões de comunicação (meios verbais e/ou não-verbais) necessários à transmissão do conhecimento seleccionado aos receptores da tradução.

Assumimos que o *feedback* dos especialistas deverá, também, levar em linha de conta os recursos empregues durante o processo de localização e os factores que

¹³² Recordamos aqui que, para nós, numa localização, o original assume características de um objecto fluido, sujeito a modificações, deixando de existir um autor identificável como sujeito individual: num projecto de localização de ontologias, o autor do conteúdo (da representação conceptual) é um sujeito corporativo, o mesmo que desencadeia e, em última instância, valida a versão final.

orientaram a selecção dos equivalentes. Para tal, interessa que esses elementos – fontes, ferramentas de tradução utilizadas, técnicas de tradução - sejam devidamente identificados e o seu uso ponderado. Esta tarefa interessa também ao próprio ambiente / contexto da rede, uma vez que é conducente à participação activa dos utilizadores, o que representa, normalmente, uma das preocupações e desafios colocados à concepção, desenvolvimento e manutenção deste tipo de espaço colaborativo.

Este *feedback* pode ser influenciado por um conjunto de processos que auxiliam no processo de localização, como o uso de metadados para capturar as decisões de tradução; a identificação/inclusão de informação sobre as fontes e recursos de conhecimento utilizados; as ferramentas de tradução e outros elementos relevantes, processos cuja sequência se pode tornar um auxiliar importante para a existência de alguma coerência nas decisões a tomar na selecção dos equivalentes e na sua validação.

Este processo de escolha e validação pode ser revisto e evoluir com o sistema, ou ser acrescentado de outros elementos que, por um lado, reforcem a percepção dos actores da rede quanto aos elementos a ter em conta no processo de localização da representação conceptual e, por outro, promovam o seu envolvimento no processo de partilha e construção do conhecimento.

O contexto da rede colaborativa¹³³, por um lado, e a simultaneidade entre conceptualização e localização, por outro, colocam-nos perante uma característica do

¹³³ Segundo Pereira (2010: 24), a abordagem proposta para partilha de informação e conhecimento no contexto de redes colaborativas deve ser de base socio-semântica pois os aspectos sociais e humanos são centrais na construção de significado partilhado. Como explicita Pereira (2010: 24), o conceito de Web Socio-semântica foi desenvolvido por Zackland (2003) e Cahier *et al.* (2003) com o objectivo de complementar a Web Semântica formal, adicionando uma abordagem pragmática. Em vez de depender inteiramente de semântica automatizada com processamento e inferência de ontologias formais, são os utilizadores (humanos) que constroem colaborativamente semântica, auxiliados por sistemas de informação socio-semânticos.

Como descreve a autora (*idem, ibidem*), incorporando a noção de socio-semântica nas arquitecturas de informação, os utilizadores do sistema são transformados numa parte crítica do seu projecto. A Web Socio-semântica visa, assim, apoiar comunidades que necessitam de colectivamente ilustrar, de forma continua, uma parte crucial do seu conhecimento, especialmente da estrutura semântica específica dos objectos de negócio e trabalho colectivo da comunidade. O significado partilhado é construído dentro de uma actividade colectiva pelos membros da rede colaborativa que assumem tanto o papel de utilizadores como “co-designers” desta cultura socio-semântica. Tendo por base a visão de Zacklad (2005), que a autora partilha, os utilizadores não são só consumidores dos recursos semânticos projectados externamente, mas também utilizadores e criadores de modo construtivo dos seus recursos semânticos. Em situações onde a construção de artefactos com significado partilhado deve ser feita de forma cooperativa, tal como defendido por Cahier *et al.* (2005), esta é, na perspectiva de Pereira (*idem, ibidem*) a abordagem adequada.

processamento da informação que interessa ter em conta, já que este processo conta com o envolvimento activo dos diversos actores em termos cognitivos, sociais e institucionais. No fundo, estes assumem um papel duplo, consubstanciado na sua experiência e conhecimento do domínio; no conhecimento que possuem do uso da terminologia específica ao domínio; e no seu contexto socioprofissional e cultural, papel representado no modelo apresentado em seguida.

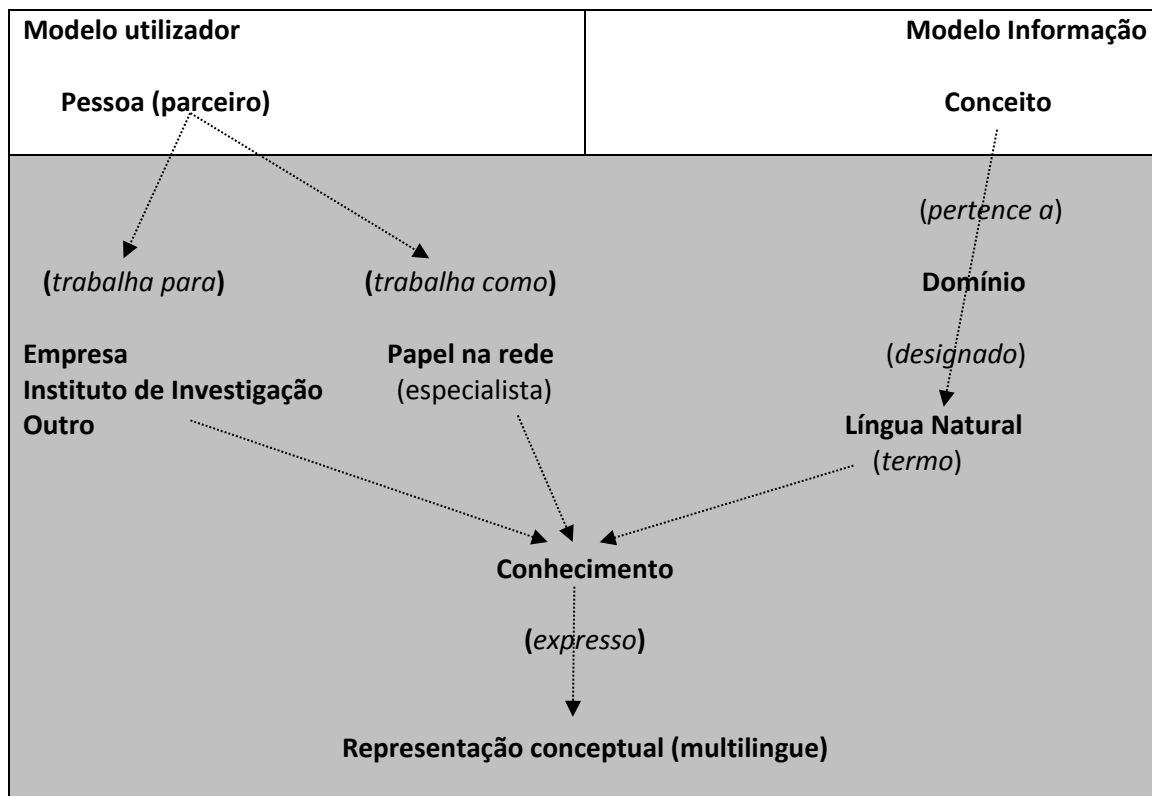


Fig. 15 – Papel do especialista

A importância deste papel¹³⁴ é destacada por Dourgnon-Hanoune *et al.* (2006) para quem a presença do especialista¹³⁵ não só permite que o conhecimento produzido seja

¹³⁴ Este papel activo e decisivo para a estruturação do conhecimento do domínio, estende-se cada vez mais a áreas da tradução (e a localização) que, como afirma O'Hagan (2009: 97), estão também elas, cada vez mais, sob o controlo crescente dos utilizadores, à medida que as comunicações electrónicas globais "*leverage the collective intelligence of users or the Internet crowd in a completely new way*".

¹³⁵ Quando consideramos o papel do especialista enquanto repositório de conhecimento especializado organizado, percebemos que é natural que cada especialista possua diferentes níveis de especialização e/ou de capacidade para comunicar sobre o seu conhecimento do domínio. Se por um lado isto é problemático, por poder gerar diferentes conceptualizações, por outro lado confere ao processo de conceptualização elementos mais ricos e mais interessantes para a negociação e a obtenção de consenso.

validado, como também identifica os termos a serem usados de modo a tornar a conceptualização da ontologia mais compreensível.

Consideramos, no entanto, e como explicitámos antes, que se a base de trabalho extralinguística, com recurso ao especialista, é um elemento que garante a partilha, reutilização e, neste caso, a transferência interlinguística do conhecimento, esta base não pode ignorar a perspectiva linguística e o recurso às fontes de conhecimento do domínio, sob pena de não se considerarem todos os elementos possíveis para a conceptualização, o que pode afectar a relevância, amplitude e profundidade do conhecimento representado, apesar de, em princípio, respeitar os propósitos de consenso, coerência e ausência de ambiguidade.

O conhecimento do domínio pelos especialistas pode, em alguns casos, e graças também ao uso de ferramentas e processos colaborativos, compensar a falta de treino formal em tradução e localização. Esta perspectiva é vista como necessária por diferentes autores, como Fu *et al.* (2009), para quem as limitações existentes ao nível dos recursos podem ter impacto no resultado do processo de selecção das traduções, a que acresce o fato de que os especialistas, *“at times, can have the expertise that is not obtained by the system, and should influence the translation selection process when necessary”*.

Os especialistas, ao validarem os resultados e ao fazerem uma escolha tradutiva, vão para além da simples análise do resultado da tradução semiautomática – a sua validação é feita com base num *background* de conhecimento do domínio considerável e estável, que fornece a capacidade de decisão quanto à escolha dos equivalentes mais significativos e auxilia no processo de análise e desambiguação dos resultados obtidos a partir da reutilização de termos provenientes de recursos de conhecimento e, eventualmente, da tradução destes.

Este conhecimento é sobretudo relevante perante a existência de mais do que um equivalente, podendo o especialista optar pelo mais apropriado, de acordo com os propósitos definidos e com as perspectivas que possui dos utilizadores finais (nos quais se inclui) e do uso da informação no seu contexto cultural local. Pode, também, propor equivalentes que não tenham sido fornecidos por qualquer tradução ou que não estejam referenciados nos restantes recursos, podendo esta interacção conduzir a uma análise e uma evolução mais rápida do processo de especificação multilingue da ontologia.

Como descreve o documento ISO/TC 37/SC 1N 326 (2006: 04), a experiência demonstra que a consciência crítica do utilizador estará mais desperta quando um termo lhe é proposto na sua língua materna do que quando adopta um termo em língua estrangeira sem o examinar. Esta consciência estará ainda mais presente se lhe for proposto que substitua, caso necessário, um termo na sua língua por outro que considere mais apropriado. Os seus autores consideram, assim, esta consciência crítica como sendo uma manifestação bastante legítima da competência linguística do *“speaker who provides terminology judgement based on criteria personal to him”*.

Este processo de localização coloca-nos, até certo ponto, perante um novo modelo de processar a informação, que é preciso observar e entender, uma vez que inclui estados cognitivos, atitudes, objectivos e emoções, dado o papel central desempenhado pelos actores da rede envolvidos no processo de construção da ontologia e na estruturação e disseminação do conhecimento, no fundo, em todo o processo comunicativo especializado¹³⁶.

A participação activa dos especialistas no processo de localização é, assim, potenciada pelas aplicações da *Web 2.0* e pela tendência cada vez mais presente da *“cultura de participação”* (Jenkins, 2006). A colaboração permitida pelos meios tecnológicos e plataformas colaborativas, de que são exemplo o espaço e as funções de partilha de aplicações como o *CmapTools*, permite a criação de grupos de trabalho e distribuição de tarefas de modo a desenvolver um projecto em moldes idênticos aos projectos de localização de *software*¹³⁷.

Interessa, finalmente, considerar que este processo de localização, participado e partilhado, é diferente do da localização de *software* efectuado por um localizador na medida em que aquele, ao contrário dos especialistas da rede:

¹³⁶ Estamos próximo do que Holden (2002) chamou de *tradução interactiva*, um modo de trabalho intercultural em que os participantes negociam *“common meanings and common understandings and learn how to be able to work in multicultural teams”*.

¹³⁷ Neste contexto, o conceito de conteúdo produzido pelo utilizador está em consonância com o conceito actual de Internet, ela própria considerada como sendo gerada, em determinados níveis, pelo utilizador. Como afirma O’Hagan (2009: 123): *“The revolution age has come to empower mere “users” in an unprecedented manner whereby enabling them to contribute their genre-knowledge and to become part of a community of practice of a new kind”*.

- 1) Não está envolvido na construção do artefacto;
- 2) Não conceptualiza – apenas localiza;
- 3) Não tem interesse directo no produto/artefacto – apenas o cliente final tem esse interesse;
- 4) Pode não conhecer os pressupostos do seu desenvolvimento ou do seu uso;
- 5) Pode nunca usar os resultados das suas decisões;
- 6) Não decide nem influencia a necessidade de revisão e alteração do produto final.

De facto, o processo de localização de ontologias corresponde, ele próprio, a um processo de descoberta, elicitación e representação de conhecimento, logo diferente do da localização de *software* ou páginas *web*, onde este fenómeno ocorre apenas ocasionalmente. Os actores da rede funcionam como especialistas e colaboram, assim, nas diferentes etapas, nomeadamente:

- 1) Na escolha e validação dos recursos de conhecimento internacionais e locais;
- 2) No processo de negociação e representação do conhecimento de domínio;
- 3) No processo de revisão e validação da estrutura do conhecimento e dos equivalentes nas diferentes línguas.

Por outro lado, o anonimato do “especialista localizador” no seio da rede colaborativa só é referente ao uso do resultado – no processo de conceptualização e de localização em si, todos os intervenientes são bem reconhecidos, tal como o papel que desempenham.

CAPÍTULO IV

4. Especificação multilingue de ontologias: experimentação

Neste capítulo apresentamos a experiência desenvolvida com o fim de implementar e validar o método proposto para a especificação multilingue de ontologias, realizada com base num caso de estudo. A experiência, que descrevemos ao longo do capítulo, foi realizada no âmbito do projecto H-Know - *Advanced Infrastructure for Knowledge Based Services for Buildings Restoring*¹³⁸ (H-Know), e consistiu na aplicação do método na especificação multilingue de uma ontologia do domínio da Reabilitação, construída de modo partilhado, no âmbito da rede colaborativa do projecto H-Know.

O capítulo descreve e contextualiza o caso de estudo, bem como as diferentes fases seguidas na implementação da experiência, apresenta e analisa os principais resultados decorrentes da experiência realizada, tendo em conta também as hipóteses que colocamos e que orientaram este estudo, e considera, finalmente, as limitações decorrentes da aplicação do método.

De modo a implementar o método e a recolher os resultados da experiência para análise, agimos, enquanto investigador participante no projecto H-Know, como observadores participantes no desenvolvimento do caso de estudo, assumindo, para além do papel de observador, os papéis de terminólogo e tradutor/localizador.

4.1. Caso de estudo – especificação multilingue da ontologia H-Know

Para o desenvolvimento desta dissertação e a experimentação do método proposto, seleccionamos como caso de estudo o projecto europeu H-Know - *Advanced Infrastructure for Knowledge Based Services for Buildings Restoring*. Este projecto, que envolveu parceiros de cinco países europeus¹³⁹, teve como objectivo a construção de uma plataforma colaborativa que permitisse o acesso ao conhecimento do domínio da reabilitação, restauro e manutenção do património edificado, sendo dirigido fundamentalmente às Pequenas e

¹³⁸ Projecto Europeu do 7º Programa Quadro, que se iniciou em Janeiro de 2009.

¹³⁹ Portugal, Espanha, França, Itália e Alemanha.

Médias Empresas (PME) e aos Institutos de Investigação e Desenvolvimento (I&D) daquele sector.

O projecto H-Know, tal como retratado na figura seguinte, preconizava a criação de um ambiente tecnológico específico que promovesse e fomentasse a partilha e reutilização do conhecimento especializado das empresas e dos especialistas do domínio de modo interactivo e colaborativo.

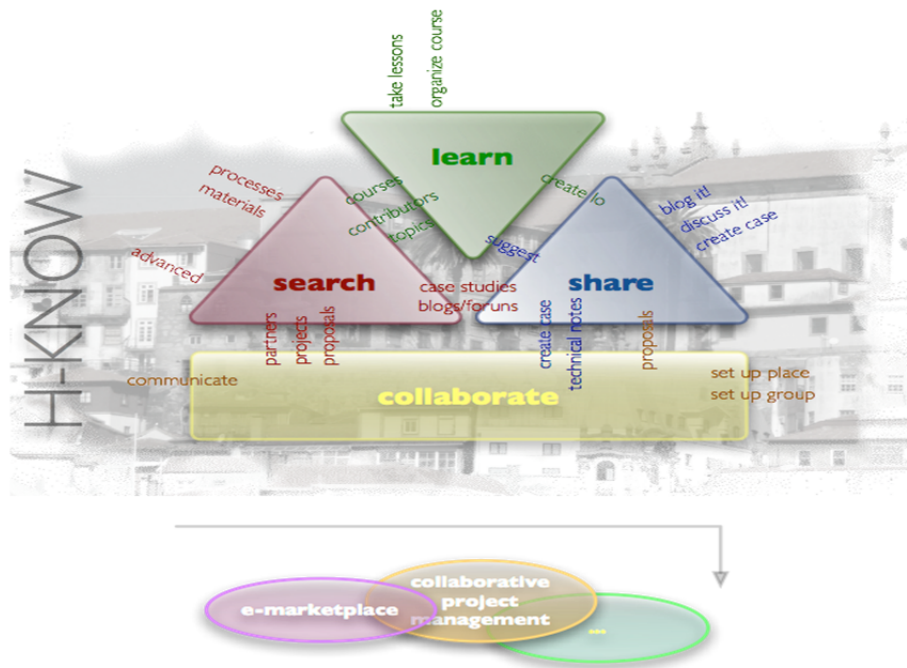


Fig. 16 – Conceito de alto nível do sistema H-Know

Como referimos anteriormente, assiste-se, hoje, a um aumento de qualidade nos processos de interacção e partilha dos recursos de conhecimento das organizações, através, por um lado, de um maior envolvimento dos diferentes interlocutores em formas eficazes e inovadoras de colaboração, como é o caso da criação de redes colaborativas e, por outro, do desenvolvimento de sistemas de gestão da informação e do conhecimento mais robustos, como os sistemas de gestão do conhecimento baseados em ontologias.

O projecto H-Know procurou, neste contexto, desenvolver uma abordagem inovadora para apoiar a criação de uma rede colaborativa na área da reabilitação do

património e de edifícios antigos¹⁴⁰, através da criação de uma plataforma colaborativa, cuja arquitectura se representa abaixo, que proporcionasse aos actores envolvidos no projecto e aos futuros utilizadores da plataforma um acesso avançado e sistemático ao conhecimento sobre processos, materiais e actores neste domínio específico da Indústria da Construção (IC). Para alcançar esses objectivos, a investigação desenvolvida abrangeu as seguintes vertentes:

- desenvolvimento de metodologias para a criação de redes virtuais de PME e Institutos de I&D, apoiadas por um conjunto de modelos e de métodos de criação de redes e espaços colaborativos. Estes têm em conta os aspectos organizacionais, culturais e técnicos e os procedimentos e modelos de estruturação e gestão do conhecimento especificamente adaptados às PME do sector;
- desenvolvimento de plataformas de apoio, albergando um conjunto de serviços para a gestão das interacções sociais, gestão do conhecimento e de formação em *e-learning*;
- desenvolvimento de serviços de apoio à gestão de conhecimento que resultassem na partilha eficiente do conhecimento do domínio no seio da rede colaborativa (H-Know, 2010)¹⁴¹.

¹⁴⁰ Pereira (2010: 38), partindo de abordagem sócio-semântica, entende a conceptualização das redes colaborativas como redes epistémicas, onde são representados os relacionamentos entre os actores da rede e as estruturas conceptuais. Tendo por base Roth (2006), a autora considera uma rede epistémica como sendo composta por três redes: *“uma rede social (envolvendo ligações entre indivíduos, neste caso representando organizações), uma rede semântica (com ligações entre conceitos e sistemas de conceitos) e uma rede socio semântica (a qual liga indivíduos a sistemas de conceitos). No final todos os parceiros pertencentes a rede colaborativa estarão ligados a um sistema de conceitos da mesma forma, ou seja, a rede semântica que representa a conceptualização partilhada”*. A rede colaborativa H-Know assume, a nosso ver, e tendo em conta a definição anterior, características de rede epistémica.

¹⁴¹ <http://www.h-know.eu/>

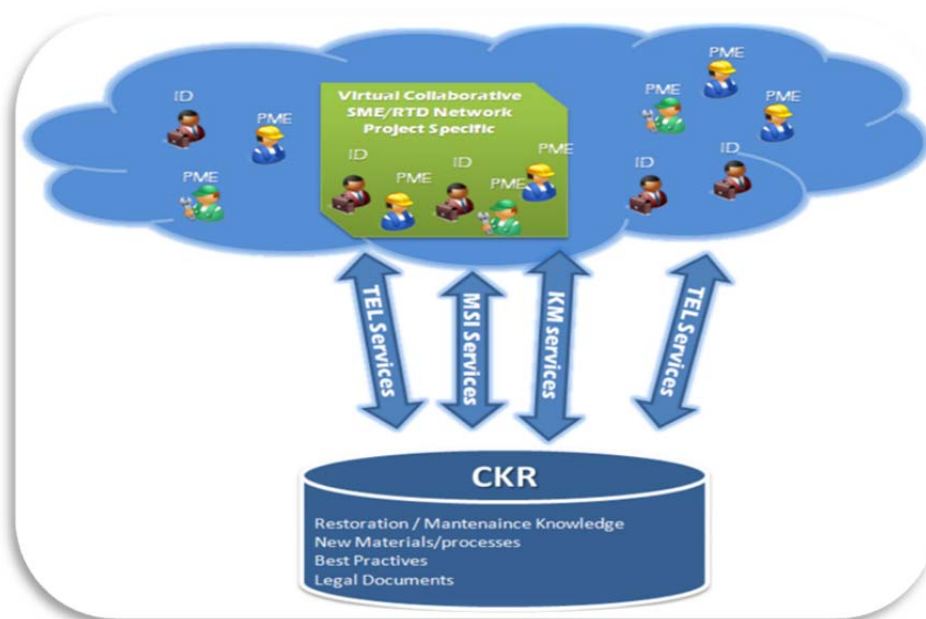


Fig. 17 – Arquitectura do sistema H-Know

Este projecto procurou, assim, através do desenvolvimento de um conjunto de métodos e ferramentas dedicados à gestão do conhecimento, à formação específica e à interação social, implementar soluções inovadoras, centradas na partilha e reutilização de conhecimento no seio de uma comunidade colaborativa constituída por PME, Associações e Institutos de I&D da área da reabilitação¹⁴².

Centrou, por isso, parte da sua atenção nos aspectos relacionados com a construção social de uma conceptualização partilhada do conhecimento, nos acordos sociais feitos sobre uma conceptualização do domínio, nos diferentes tipos de actores envolvidos no processo e nas metodologias e ferramentas disponíveis para apoiar a construção da conceptualização do domínio e a sua localização, bem como na identificação dos recursos e fontes de informação disponíveis para apoiar os processos de conceptualização e localização.

Seguiu-se, deste modo, uma abordagem à gestão do conhecimento baseada em ontologias com o intuito de criar uma ontologia de domínio, desenvolvida de modo colaborativo e cujos objectivos se centraram em:

¹⁴² A comunidade que formava a rede colaborativa H-Know pode ser descrita, de modo breve, como uma entidade sociocultural complexa e especializada, de carácter multicultural e multilingue, com objectivos de colaboração e partilha, composta por um conjunto de intervenientes especialistas cujo saber é diversificado e transdisciplinar.

1. *develop an infrastructure to efficiently and effectively organize, classify and retrieve information and knowledge,*
2. *supply H-Know users with a common ground for a shared understanding of terms and concepts when engaging in the VBE/VCN¹⁴³ activities, and*
3. *develop a tool for interoperability when proprietary databases and knowledge bases need to communicate with the H-Know system. (Soeiro, 2010: 03)*

A abordagem seguida no âmbito do projecto para a construção da ontologia de domínio considerou dois níveis e etapas distintas. A primeira etapa centrou-se no desenvolvimento de uma ontologia de alto nível – a *H-Know Global Ontology* (H-KnowG), criada, num primeiro momento, com recurso à língua inglesa, e, posteriormente, localizada para as línguas dos diferentes parceiros do projecto. Numa segunda etapa, proceder-se-ia à construção de ontologias locais – *H-Know Local Ontologies* (H-KnowL), que seriam posteriormente alinhadas e integradas com a ontologia global, tal como se descreve na figura seguinte.

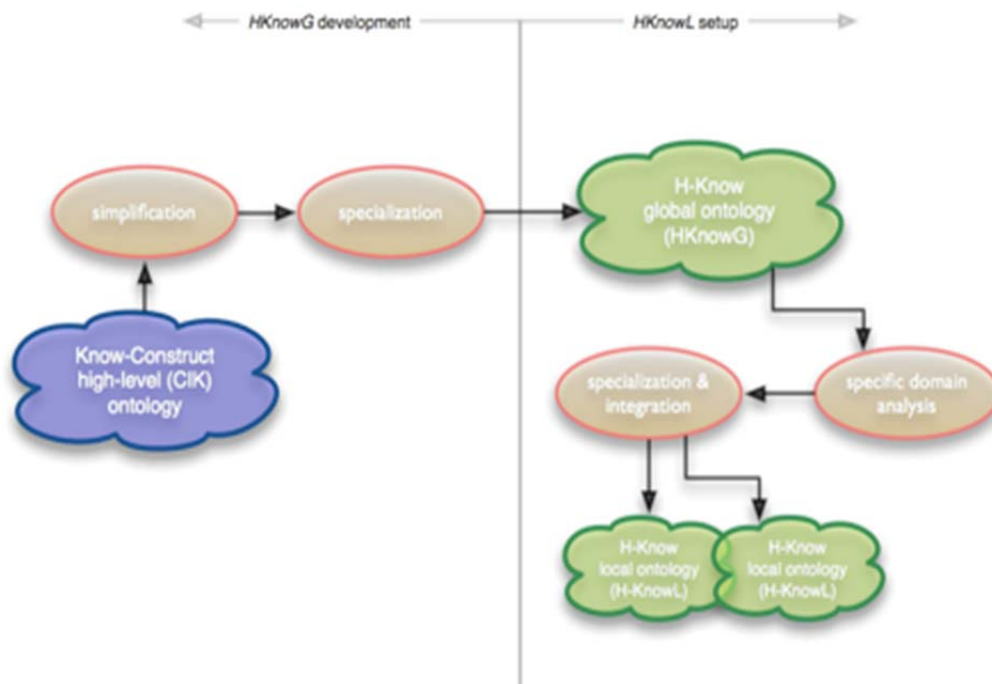


Fig. 18 – Processo de construção da ontologia H-Know:
da H-KnowG (global) às ontologias H-KnowL (locais)

¹⁴³ VBE – Virtual Breeding Environment; VCN – Virtual Collaborative Network.

A decisão de seguir esta metodologia foi tomada após análise cuidada dos requisitos apresentados pelos actores do projecto, de modo a responder às necessidades profissionais, culturais e linguísticas particulares de cada parceiro H-Know, bem como ao contexto social e profissional em que a(s) ontologia(s) e a plataforma H-Know seriam utilizadas.

Por outro lado, as duas etapas correspondem a necessidades e objectivos diferentes da rede H-Know: a primeira, de estruturação e representação da informação do domínio de mais alto nível e de obtenção de consenso entre os especialistas dos diferentes países; a segunda, de estabelecimento de negócios e partilha de conhecimento a um nível mais local, em que, previsivelmente, será necessário o uso de informação e conhecimento mais específico. Esta decisão de optar pela construção de mais do que uma ontologia representando níveis distintos, seguiu abordagens testadas em projectos anteriores do domínio da IC, como o *Know-Construct*¹⁴⁴, e as recomendações da norma ISO 12006-2¹⁴⁵.

A solução adoptada pretendeu, assim, desenvolver um sistema de gestão do conhecimento construído com base em “*distributed ontologies, locally managed and centrally integrated into H-KnowG, with intention to adopt in the future the usage of a global ontology*” (H-Know, 2010)¹⁴⁶. Em termos tecnológicos foi tomada a decisão de implementar e fazer a manutenção da ontologia utilizando as aplicações Protégé¹⁴⁷ e OntoWiki¹⁴⁸.

4.2. Implementação do método no contexto do caso de estudo

Como resulta dos objectivos do projecto H-Know e da tipologia multilingue da sua rede de parceiros, tanto o conhecimento do domínio a disponibilizar, como a representação

¹⁴⁴ <http://www.know-construct.com/index.htm>

¹⁴⁵ Como afirma Ekholm (2005: 275), o sector da construção assume, tradicionalmente, um carácter nacional e regional. Refere, no entanto, o desenvolvimento crescente de normas que procuram contribuir para a normalização da gestão de informação no sector, de que destaca duas: “*there are two major international candidates for core ontologies common to the sector, ISO 12006-2:2001, Building construction - Organization of information about construction works – Part 2: Framework for classification of information (ISO 2002), and Industry Foundation Classes, IFC, developed by the International Alliance for Interoperability, IAI (IAI 2000)*”.

¹⁴⁶ H-Know D1.3 System Concept. <http://h-know.eu/project-overview/wp-deliverables>

¹⁴⁷ <http://protege.stanford.edu/download/ontologies.html>

¹⁴⁸ <http://ontowiki.net/Projects/OntoWiki>

desse mesmo conhecimento em termos ontológicos, não poderiam ficar restringidos ao uso de uma língua em particular, pelo que o recurso a uma representação multilingue se tornou uma necessidade evidente no processo de construção da ontologia H-Know, tendo a localização surgido como o processo natural a seguir para desenvolver a sua especificação multilingue.

Esta situação suscitou questões no âmbito do projecto quanto à metodologia a seguir, o tipo de aplicações de auxílio à localização a utilizar e quanto ao envolvimento dos especialistas, grande parte dos quais seria confrontado, pela primeira vez, com o próprio conceito de localização.

Estas questões e a procura de uma proposta para a sua resolução resultaram no desenvolvimento e na implementação do método antes descrito no âmbito deste projecto, como forma de auxiliar o processo de localização e da própria conceptualização da ontologia H-Know.

Procurámos, assim, através da aplicação deste método ao desenvolvimento do caso de estudo em questão criar as condições para a integração e uso do multilinguismo no processo de conceptualização, no contexto de uma rede colaborativa e de um ambiente de partilha do conhecimento. Paralelamente, procurámos identificar estratégias de modo a minorar as dificuldades que a presença de mais do que uma língua na fase de conceptualização pudesse causar e a permitir que o processo de localização, desenvolvido no momento subsequente à conceptualização, conduzisse à obtenção de maior *feedback* por parte dos especialistas a partir do seu trabalho sobre as representações semiformais de conhecimento do domínio.

A implementação do método teve em conta a diversidade do grupo de utilizadores e dos seus requisitos, os recursos existentes no momento da sua concepção e implementação, os objectivos do projecto e do sistema a desenvolver e as condicionantes colocadas pela integração dos resultados da localização numa aplicação de construção de ontologias já existente, neste caso a Protégé. Nesse sentido, e para a implementação e experimentação do método, foi considerado estarem reunidas as seguintes condições:

1. a rede colaborativa está formada, é multilingue e os seus objectivos, missão e prazos estão definidos, sendo aceites por todos os seus membros;

2. os parceiros, nomeadamente aqueles mais directamente ligados ao domínio da IC, serão os interlocutores envolvidos no processo de conceptualização do conhecimento do domínio;
3. cada parceiro é visto como um especialista e participa activamente na conceptualização da ontologia do domínio e, simultaneamente, no processo de localização, de acordo com os papéis, os propósitos e a calendarização definidos.

Como referimos já, um outro aspecto previamente definido era o do domínio de especialidade sobre o qual incidiria o projecto H-Know: o da reabilitação, domínio que esteve na base do desenvolvimento de todos os cenários e processos de recolha, gestão e partilha da informação e conhecimento desenhados e desenvolvidos ao longo do projecto e da experimentação do método.

Tendo por base o contexto e os objectivos a atingir no âmbito do projecto H-Know, procurámos implementar e testar o método proposto para a especificação multilingue de ontologias através da condução de uma experiência no âmbito da construção da ontologia H-Know. Para proceder à experimentação, recolha e análise dos resultados deste caso de estudo, participámos, como afirmámos antes, enquanto investigadores do projecto, como observadores participantes no projecto de investigação e desenvolvimento tecnológico, assumindo, para além do papel de observador, os papéis do terminólogo e, em parte, do tradutor/localizador.

Em função dos componentes e passos propostos no método, descritos no capítulo anterior, e do caso de estudo definido, concebemos um plano para experimentar o método dividido em seis fases, sendo cada uma destas fases composta por diferentes passos, actividades e objectivos que, em seguida, enumeramos e descrevemos, de forma breve:

- Fase 1 – Selecção de recursos do domínio de especialidade e criação do ambiente de localização
- Fase 2 – Definição dos participantes e da amostra
- Fase 3 – Criação do espaço de conceptualização e localização
- Fase 4 – Experimentação do método
- Fase 5 – Análise de resultados
- Fase 6 – Reestruturação do método

A primeira fase refere-se ao processo de recolha e selecção de recursos do domínio para a constituição do *corpus* de especialidade. Nesta procedemos, primeiro, à caracterização e delimitação do domínio e, em seguida, à compilação de recursos para a construção do *corpus*.

Na segunda fase, procedemos à definição dos participantes na experiência. Para tal, procedemos a selecção de um grupo de especialistas do domínio representativo das diferentes línguas de trabalho a partir dos actores da rede colaborativa. Paralelamente, procedemos à definição da amostra dos mapas conceptuais que constituíram o nosso objecto de análise, a partir do conjunto dos mapas conceptuais resultantes da conceptualização inicial, desenvolvida em inglês.

Na terceira fase, procedemos à criação do espaço de conceptualização e localização, com recurso à aplicação CmapTools, de modo a dotar os especialistas de um espaço e de um fluxo de trabalho que apoiasse o processo de enriquecimento multilingue de cada representação conceptual. Definimos, também neste passo, após um processo de análise do seu funcionamento, quais as aplicações *web* a utilizar para apoiar a realização do processo de localização no contexto descrito.

Definidos os recursos, o ambiente de trabalho, os especialistas e a amostra, iniciámos a experimentação do método, na quarta fase. De modo a apoiar e a guiar o especialista durante esta fase de desenvolvimento dos processos de análise conceptual e de localização, elaborámos um guião com informações sobre o tipo de acções a empreender e os recursos a utilizar. Para explicitar as dúvidas sobre a realização da experiência ou sobre as aplicações a utilizar, por um lado, e, por outro, para observar as dificuldades e problemas sentidos pelos especialistas na localização dos termos do domínio e auxiliar na sua resolução, acompanhámos, na sua totalidade, o desenrolar desta fase.

A quinta fase foi dedicada à análise dos resultados obtidos com a experimentação do método, a partir da observação das decisões de (re)conceptualização e de localização feitas com base nas representações conceptuais iniciais. Esta análise foi consubstanciada pela realização de um inquérito aos especialistas, através de uma entrevista semiestruturada, que procurou obter a opinião dos mesmos sobre o uso do método proposto e inquirir sobre a sua percepção quanto ao papel desempenhado na escolha e definição dos equivalentes, bem como sobre as suas dificuldades e limitações para desenvolver esta actividade. Uma vez

que o inquérito procurou obter a opinião dos especialistas sobre a aplicação do método nas suas diferentes fases, actividades e objectivos, parte das respostas obtidas serão transcritas e consideradas em diferentes momentos deste capítulo, nomeadamente nos pontos 4.2.1.2 e 4.2.3.1, e não apenas no ponto 4.2.5.2, dedicado à análise dos resultados do inquérito.

Finalmente, na sexta fase, procurámos integrar no método os contributos resultantes da análise dos resultados e do *feedback* dos especialistas, de modo a obter uma melhoria no fluxo de trabalho proposto e na sua implementação e desenvolvimento futuro.

Descrevemos, em seguida, os passos e actividades desenvolvidos em cada fase da experiência, bem como os resultados obtidos.

4.2.1. Fase 1 – Selecção de recursos do domínio e criação do ambiente de localização

Um elemento nuclear aos processos de conceptualização e localização é, tal como apontado no ponto 3.3.1, o domínio de especialidade, dependendo a sua conceptualização, em grande parte, dos propósitos e das actividades a desenvolver, bem como de uma delimitação concreta do domínio a analisar e a representar. Esta delimitação é, também ela, uma acção intencional, regida por necessidades e objectivos específicos pré-determinados no seio da rede¹⁴⁹.

A primeira etapa para a recolha e estruturação da informação e desenvolvimento do sistema conceptual consistiu, assim, na identificação e delimitação do domínio e respectivos subdomínios de saber. Para tal procedemos a uma análise e a uma caracterização do domínio.

¹⁴⁹ Para Rondeau (1984: 12), a delimitação pode conduzir ao estabelecimento de fronteiras arbitrárias e artificiais, dado que existem, como descreve, “*d’une part, des domaines qui se chevauchent, d’autre part, des sous-domaines qui peuvent appartenir à plus d’un domaine, (...), et enfin des domaines entiers qui, par leurs applications, peuvent se retrouver dans bon nombre d’autres*”. (Rondeau, 1984: 12)

4.2.1.1. Domínio de especialidade: caracterização e delimitação

A delimitação clara do domínio de especialidade é, como descrevemos ao longo do ponto 3.3, de grande importância para o processo de conceptualização. A complexidade inerente a este processo e à representação do conhecimento de um domínio, anteriormente identificada e descrita, advém também das características do próprio domínio, da sua maior ou menor multidimensionalidade e das relações de interdisciplinaridade que estabelece com outros domínios¹⁵⁰.

O domínio de especialidade delimitado neste trabalho foi o da *reabilitação*¹⁵¹, que podemos considerar como um subdomínio do domínio da IC. Assim, em seguida, caracterizaremos primeiro o domínio da IC, destacando nele as características mais genéricas e transversais, e, posteriormente, o subdomínio da reabilitação.

A IC caracteriza-se pela sua actividade cíclica, rotativa e, até certo ponto, nómada. É, no seu dia-a-dia, uma indústria fortemente influenciada por um conjunto de factores exógenos como a evolução da conjuntura económica, do investimento público, da procura, das condições climatéricas, do recrutamento de pessoal qualificado, dos materiais, das perspectivas de vendas, das taxas de juro, entre outros condicionalismos micro e macro económicos nacionais, europeus e mundiais.

É, por isso, um domínio que ocupa um plano intermédio entre as duas categorias de domínios definidas por Bergenholtz e Tarp (1995:26), dado que neste domínio um conjunto de conceitos são comuns e transversais, como, por exemplo, os referentes a produtos, existindo outros conceitos que são dependentes da cultura, como é o caso de alguns referentes aos processos construtivos. É, para além disso, um domínio fortemente

¹⁵⁰ Autores como Bergenholtz e Tarp (1995:26) consideram fundamental a distinção entre domínios culturalmente dependentes e domínios culturalmente independentes (i.e. que são ou não afectados por especificidades locais dos diferentes países). Os domínios da segunda categoria, como a aeronáutica, apresentam, segundo os autores, termos e conceitos “universais” e “concretos” por oposição aos termos e conceitos “únicos” e “abstractos” dos domínios de especialidade da primeira categoria, como o jurídico. Esta distinção tem características análogas à proposta por Espinoza *et al.* (2009), que referimos no ponto 2.3.1.

¹⁵¹ Este domínio é também designado por: *reabilitação do património, reabilitação do património edificado, reabilitação do património construído, reabilitação de edifícios (antigos)*. Adoptaremos, ao longo deste trabalho a designação mais corrente e mais genérica de *domínio da reabilitação*.

condicionado pela legislação de cada país, que influencia a própria estrutura organizacional do processo construtivo¹⁵².

A IC envolve, tradicionalmente, um elevado número de processos de *design* e de produção, pelo que o conhecimento existente nesse domínio pode ser descrito, de modo breve, como sendo disperso, diverso e fragmentado, em resultado do seu carácter polimórfico e da quantidade de actores, normas e instituições que intervêm no desenrolar de todo o processo construtivo, conhecimento que provém e é composto por disciplinas de diversas áreas disciplinares, que se sobrepõem, integram e se complementam.

O próprio processo construtivo é visto e descrito de modo diferente pelos vários autores, tal como se percebe pelas figuras seguintes, mesmo quando os autores são originários do mesmo país e da mesma escola de pensamento, como é o caso dos processos descritos nas duas últimas figuras, da autoria de professores da mesma instituição universitária, o que espelha a complexidade da área.

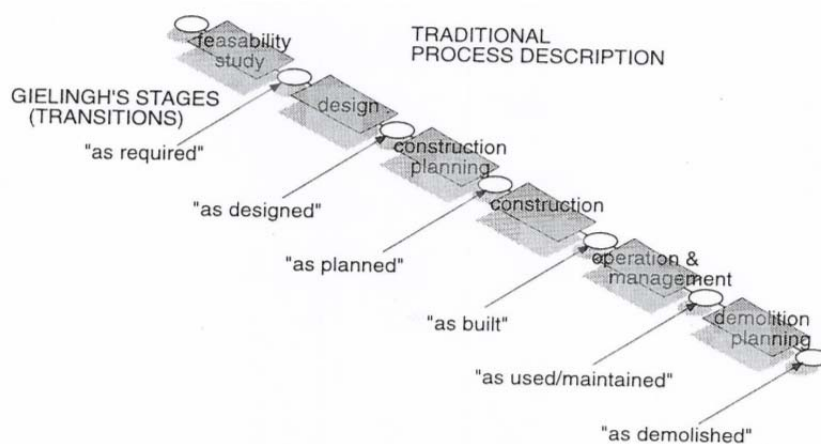


Fig. 19 – Processo construtivo I (*building lifecycle*) (Eastman, 2008)

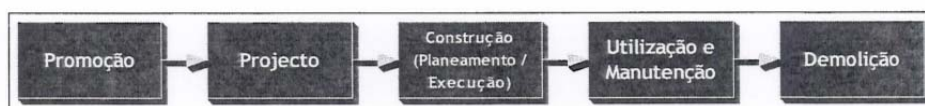


Fig. 15 – Processo construtivo: esquema simplificado (Poças Martins 2006)

Fig. 20 – Processo construtivo II (Poças Martins, 2006)

¹⁵² Ver Anexo 2.

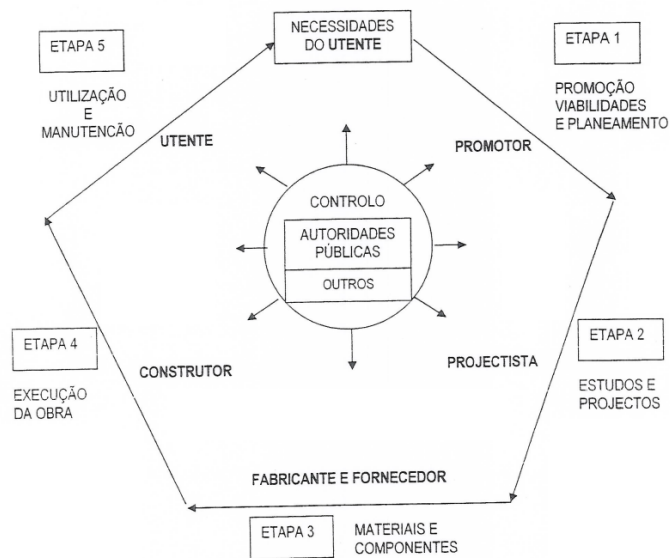


Fig. 21 – Processo construtivo III (Hipólito de Sousa, 2001)

Por outro lado, o ciclo de inovação tecnológica neste sector é, tipicamente, tão fragmentado como a indústria em si mesmo. Não sendo necessariamente um problema, uma vez que a fragmentação pode ser vista como conduzindo à flexibilidade, pode, no entanto, conduzir a ruídos na comunicação entre os diferentes interlocutores e à falta de interoperabilidade entre os seus sistemas tecnológicos e a entendimentos particulares das actividades, mecanismos e resultados inerentes ao processo de construção.

Esta diversidade e fragmentação obrigam a um cuidado especial na estruturação e representação do conhecimento existente, de modo a garantir a transparência das representações conceptuais obtidas e a auxiliar na construção de um entendimento comum¹⁵³ quanto ao uso dos termos, de modo a aumentar a sua consistência e a normalizar o seu uso.

Esta necessidade, reconhecida e transversal a vários países e projectos tem conduzido ao desenvolvimento de esforços no sentido de estruturar a informação no domínio da IC, sobretudo ao nível da construção de classificações mais ou menos

¹⁵³ Um exemplo da falta de entendimento sobre o significado dos conceitos neste domínio ressalta do uso do termo “elemento” que, na norma ISO (1994) *Classification of information in the construction industry*, é definido como “one characteristic function of a physical part of a building”, enquanto na norma ISO (1995) *Building elements using explicit shape representation*. Part 225 of ISO 10303, dedicada à modelação do produto, é descrito como “a combination of properties”.

específicas¹⁵⁴. Estas classificações e outros recursos seguem, no entanto, uma abordagem de organização e estruturação da informação que se focaliza, como pudemos observar, nas necessidades de um país ou grupo de países específicos, tendendo o seu uso, não raramente, a ser circunscrito¹⁵⁵.

Em Portugal foram já desenvolvidos projectos de estruturação da informação nesta área, como aconteceu com o projecto europeu Know-Construct¹⁵⁶, o *Cic-Net - Rede de Cooperação Estratégica entre Empresas do Processo de Construção*, entre outros projectos de investigação, havendo também um conjunto de especificações e terminologias oriundas de entidades como o Laboratório Nacional de Engenharia Civil (LNEC) que pretendem dar resposta a esta necessidade.

Verifica-se, por outro lado, a existência de sítios na Internet, como o Grupopuma¹⁵⁷, o Construlink¹⁵⁸ ou o SIKA¹⁵⁹, que apresentam a sua própria classificação multilingue do conhecimento do domínio da IC, nestes casos, adequada ao tipo e às necessidades de negócio de cada empresa¹⁶⁰, classificações que estão sobretudo ligadas à comercialização de produtos de construção no espaço europeu.

Sendo este um domínio vasto e complexo, delimitamos o nosso enfoque no subdomínio da *reabilitação*, subdomínio que herda, no conjunto dos seus processos e actores, muitas das características supradescritas que enformam a IC.

A reabilitação pode ser definida como o acto ou processo de possibilitar um uso eficiente e compatível de um edifício através de reparações, alterações ou ampliações,

¹⁵⁴ Exemplo: Na Europa, as classificações EPIC (1999) e Uniclass (1997). Nos EUA, a classificação CSI MasterFormat (2003).

¹⁵⁵ Uma classificação que seguiu uma perspectiva diferente – a da universalidade e transversalidade - foi a UNSPSC (2003), desenvolvida pela ONU, que, no entanto, tem sido pouco difundida e utilizada.

¹⁵⁶ Know-Construct - *Internet Platform for Knowledge-based Customer Needs Management and Collaboration among SMEs in Construction Industry* - COLL-CT-2004-500276 (www.know-construct.com).

¹⁵⁷ <http://www.grupopuma.com/pt/index.php>

¹⁵⁸ www.construlink.com

¹⁵⁹ www.sika.pt/home.htm

¹⁶⁰ A classificação usada nestes sítios *Web* refere-se sobretudo a produtos e materiais de construção e está localizada para diferentes línguas, dependendo dos mercados em que cada empresa opera.

preservando as características que transmitem os seus valores histórico, cultural e arquitectónico¹⁶¹.

Na reabilitação, a actividade construtiva incide principalmente no campo da acção arquitectónica e urbanística e inclui, muitas vezes em simultâneo, a prevenção da deterioração, a conservação e o restauro, a reforma, a manutenção, a ampliação e a transformação, a reconstrução e a renovação, sendo o seu objectivo o de dotar o objecto edificado ou conjunto histórico de melhores condições de habitabilidade ou de uso.

A reabilitação, sobretudo a urbana, tornou-se um elemento incontornável quer a nossa atenção se focalize nos aspectos de conservação ou defesa do património, de desenvolvimento sustentável, de ordenamento do território ou de coesão social. O conceito de reabilitação sofreu, por isso, uma grande evolução nas últimas décadas, no que respeita aos seus princípios, objectivos, âmbito de actuação, metodologia e abordagem. Parte da política de conservação do património arquitectónico, mas ultrapassa esse âmbito, em resposta aos novos desafios de natureza social, económica, ambiental e cultural.

Este domínio esteve, como referimos, no centro do desenvolvimento do projecto H-Know e da criação de uma comunidade colaborativa na área da reabilitação. Neste contexto, uma das actividades do projecto consistiu na identificação das fontes de informação relativas ao domínio da reabilitação, sobretudo as respeitantes às áreas do património, da herança cultural e da manutenção e reabilitação de edifícios antigos, de modo a estruturar a informação recolhida e a utilizá-la como base para a construção da ontologia H-Know e para a sua especificação multilingue. Este processo foi desenvolvido com a participação dos especialistas e do terminólogo.

Foi a partir desta tarefa que o nosso envolvimento com o projecto teve o seu início, constituindo este também o primeiro elemento do nosso caso de estudo e a primeira fase do método.

¹⁶¹ <http://www.portaldahabitacao.pt/pt/portal/glossario/consultarAlfabetoList.jsp>

4.2.1.2. Compilação do *corpus*

Um passo essencial para a conceptualização e a localização, na perspectiva do método que propomos e de uma metodologia de base terminológica, é a constituição de um *corpus* específico do domínio, composto por recursos representativos das diferentes línguas de trabalho, recolhidos de acordo com critérios pré-estabelecidos e passíveis de serem utilizados no processo de representação do conhecimento do domínio, pelo que este se constituiu necessariamente como o primeiro passo para a construção e aplicação de um método de apoio à especificação multilingue de ontologias.

O trabalho de representação do conhecimento de um domínio deve ser desenvolvido, idealmente, por equipas multi e transdisciplinares, nas quais o terminólogo pode funcionar, como descrito no ponto 3.4, como um mediador, sendo uma das suas tarefas a de propor uma metodologia de trabalho, que pode ser adequada aos objectivos de cada projecto ou grupo profissional. A compilação de um *corpus* constitui uma das fases dessa metodologia e dá resposta a necessidades várias que ocorrem em diferentes etapas e tarefas do trabalho a desenvolver para aceder, estruturar, negociar e representar o conhecimento de um domínio de especialidade.

No âmbito do presente caso de estudo e seguindo os princípios e critérios enunciados nos pontos 3.3.1 e seguintes, procedemos à compilação de um conjunto de recursos relevantes para o domínio, de tipo terminológico, lexical, ontológico, taxonómico e outros, que apoiasse os processos de conceptualização e de localização. A compilação destes recursos específicos teve em conta os critérios estabelecidos na tabela 1, ponto 3.3.1.1, e considerou as possibilidades e os efeitos da sua reutilização na delimitação e representação do domínio e na conceptualização e localização da ontologia, tendo a selecção sido validada pelos especialistas do domínio afectos ao projecto.

A reutilização dos recursos existentes foi definida como sendo a abordagem preferencial a seguir na construção da ontologia H-Know e como o primeiro passo para a construção da estrutura de recursos semânticos de apoio ao sistema H-Know, o qual seria utilizado na gestão do conhecimento da rede colaborativa multilingue da área da reabilitação.

O princípio de reutilização de recursos semânticos foi seguido também com o intuito de promover o uso mais alargado das fontes consideradas mais relevantes e, assim, contribuir, por um lado, para a implementação de normas no domínio da IC, em geral, e da reabilitação, em particular e, por outro, corresponder à necessidade de contextualizar e harmonizar os conceitos de uma indústria onde o conhecimento é essencialmente aplicado a nível local, de modo a evitar interpretações incorrectas e a diminuir o ruído na comunicação. Considerou, finalmente, o âmbito, propósitos e necessidades específicas do projecto, bem como o contexto sociocultural e profissional de desenvolvimento e uso da ontologia e o seu potencial público-alvo, factores que estiveram também na base do processo de adequação de cada recurso.

Sendo a IC um domínio complexo, em que o conhecimento progride rapidamente, houve, no entanto, necessidade de rever e readequar o conhecimento implementado. De facto, quando consultamos os múltiplos recursos existentes e temos em conta os seus diferentes tipos, verificamos que o problema da reutilização nem sempre é simples de resolver, uma vez que, como identifica Roche (2006: 584), uma parte das fontes de informação tem a sua origem em instituições cujo objectivo principal não é o da descrição do conhecimento, mas antes a descrição do seu uso, para além de serem construídas a partir de diferentes princípios teóricos e com propósitos distintos dos enunciados para a construção da ontologia H-Know.

A análise das fontes identificadas com a intervenção dos especialistas permitiu, no entanto, concluir que uma parte dos recursos existentes possuía princípios e estruturas comuns, uma vez que emanavam de projectos governamentais, europeus ou de associações internacionais do sector, que procuravam também contribuir para a harmonização e normalização neste domínio, como se pode verificar pelo exemplo analisado na tabela seguinte, onde se comparam as propostas de classificação de organização da informação sobre o processo construtivo – a da UNICLASS e da ISO 12006-2.

Tabelas UNICLASS		Tabelas ISO CD12006-2	Construções de dimensões relevantes (pela forma)	Construções de dimensões relevantes (função /actividade uso)	Construções complexas (por função ou actividade)	Espaços (pelo grau de enclausuramento)	Espaços (pela função ou actividade de uso)	Instalações (resultados de construção por função ou actividade de uso)	Elementos de Construção	Elementos de construção caracterizados (elementos pelo tipo de actividade)	Actividades de construção	Gestão do processo construtivo (pelo tipo de processo)	Ciclo de vida das construções	Fases das construções	Produtos de construção	Recursos complementares à construção	Agentes de construção	Informação na construção (por tipo)	Propriedades e características
A	Tipo de informação																		x
B	Disciplinas envolvidas																x		
C	Direcção e gestão											x		x					
D	Instalações		x	x		x	x												
E	Construções de dimensões relevantes	x																	
F	Espaços				x														
G	Elementos de Construção para edifícios								x										
H	Elementos de Construção para engenharia civil								x										
J	Lista de actividades de construção para edifícios									x	x								
K	Lista de actividades de construção para engenharia civil									x	x								
L	Produtos de Construção														x				
M	Recursos complementares à construção															x			
N	Propriedades e características																		x
P	Matérias																		
Q	Classificação Decimal Universal																		

Tabela 2: Exemplo da análise de recursos do domínio¹⁶²

De modo completar o nosso *corpus*, e tendo em conta o método que propomos, procedemos também à recolha de um *corpus* textual comparável, neste caso para as línguas portuguesa e inglesa. Decidimos abranger apenas as línguas portuguesa e inglesa para a construção do *corpus* comparável, uma vez que o *corpus* serviu essencialmente para efeitos

¹⁶² Baseada em Monteiro, (1998: 70).

de consulta, em complemento ao trabalho com o especialista e o corpus de recursos linguísticos e semânticos, como descrevemos a seguir. Serviu também, no nosso caso, para melhor analisar, compreender e delimitar o domínio e os seus conceitos nucleares. Por outro lado, considerámos que seria altamente complexo obter e analisar os resultados advindos de um corpus comparável que incluísse as seis línguas envolvidas no projecto, dados os limites temporais e a presença de apenas um terminólogo neste projecto. Seleccionámos, portanto, a língua de partida da ontologia de alto nível e a língua de chegada do terminólogo.

Seguimos, para esse efeito, os pontos de vista e critérios advogados no ponto 3.3.1.2. Em função desses critérios, definimos uma tipologia de textos para o domínio da reabilitação, como forma de agrupar e estruturar o *corpus* e atestar a sua representatividade e abrangência.

A tipologia de textos que em seguida propomos para o domínio da reabilitação baseia-se na observação e na análise de textos escritos em português e inglês e procura constituir uma amostra equilibrada em termos de dimensão e variação dos textos e registos passíveis de serem encontrados sobre o domínio. De modo a perceber quais os tipos de texto mais relevantes para o *corpus* e para os nossos objectivos específicos, tivemos em conta os seguintes aspectos:

1. O contexto comunicativo no qual foram produzidos, que corresponde ao contexto profissional e cultural de produção;
2. As características do público-alvo previamente identificado, sendo este composto essencialmente por especialistas do domínio que farão uso dos resultados obtidos;
3. O facto de as publicações no domínio da reabilitação serem, dada a complexidade e transversalidade da área, dirigidas sobretudo a um público com formação superior, nomeadamente engenheiros, arquitectos e donos de obra;
4. O facto de, no domínio da reabilitação, todos os processos serem condicionados por normas legais, o que torna necessário seleccionar e incluir uma amostra representativa de textos jurídicos.

Tendo em conta estes aspectos desenvolvemos, para o domínio da reabilitação, a seguinte tipologia:

Tipo de texto	Tipo de documento	Contexto de produção
Texto Científico	Teses mestrado	Ensino superior
	Teses doutoramento	Ensino superior
Texto Técnico	Monografias	Editoras especializadas
	Artigos em conferência	Conferências da especialidade
	Revistas técnicas sobre construção	Entidades do sector
	Fichas sobre patologias da construção	Associações do sector
	Dossiers de especialidade	Grupo de estudos da patologia da construção
	Revistas de especialidade	Associações do sector
Texto didáctico	Manuais universitários	Ensino superior
	Manuais de formação técnica	Associações do sector
Texto jurídico	Legislação nacional e local	Órgãos Legislativos e Câmaras Municipais
	Contrato	Empresas do sector

Tabela 3: Tipologia de textos do *corpus* do domínio da reabilitação

Atendendo aos objectivos e ao público-alvo que previsivelmente iria aceder e utilizar a plataforma, não foram considerados para a constituição deste *corpus*, depois de analisados e testados, outros tipos de texto, cujo cariz era essencialmente de divulgação publicitária ou de vulgarização, uma vez que não continham informação cuja relevância ou novidade potenciassessem e promovessem a partilha de conhecimento.¹⁶³

Com base no *corpus* compilado, e modo a identificar, delimitar e representar os termos pertencentes ao domínio da reabilitação, recorreremos a uma metodologia de base terminológica mista, de carácter iterativo que, numa primeira fase, se socorreu de uma abordagem conceptual, de base onomasiológica, assente no trabalho com o especialista, a qual que foi complementada por uma abordagem linguística, de base semasiológica, assente sobretudo na reutilização de recursos semânticos, como normas, classificações, terminologias e tesouros, e no recurso, ainda que menos sistemático, ao *corpus* textual.

Foram ainda usados outros elementos, como representações gráficas e tabelas. Esta abordagem foi seguida para apoiar o processo de conceptualização do domínio, quer na sua

¹⁶³ A composição do *corpus* textual é descrita no Apêndice F.

fase inicial, de conceptualização e representação do domínio com recurso à língua inglesa quer durante o processo de localização e de reconceptualização.

Importa, nesta fase, destacar que o uso destes recursos diminuiu bastante a necessidade de recorrermos ao *corpus* textual, uma vez que a informação sobre o domínio contida naqueles, juntamente com o conhecimento fornecido pelos especialistas, foi considerada como respondendo de modo suficiente às necessidades de informação e aos requisitos definidos para a construção da ontologia de alto nível do H-Know.

O recurso ao *corpus* textual foi, assim, feito em modo complementar, sobretudo para efeitos de consulta e verificação do uso dos termos em discurso nas diferentes línguas, e como auxílio à desambiguação, sempre que esta se verificava, processo levado a cabo pelo terminólogo. Serviu também para melhor determinar os termos nucleares ao domínio, através do recurso às metodologias de extracção terminológica e ao estabelecimento de concordâncias¹⁶⁴, como auxílio à pesquisa de definições e como meio para analisar o tipo de relações que os conceitos poderiam estabelecer entre si. O recurso ao *corpus* textual ocorreu particularmente ao longo do desenvolvimento do processo de conceptualização inicial.

O ponto de partida para esta conceptualização inicial do domínio foi o modelo de apoio à classificação da informação da IC, representado na figura seguinte, proposto na norma ISO 12006-2¹⁶⁵ - *Building construction - Organization of information about construction works - Framework for classification of information*.

¹⁶⁴ Usámos, para este efeito, a aplicação WordSmith: <http://www.lexically.net/wordsmith/index.html>.

¹⁶⁵ Teve a sua origem no comité *Building construction, Subcommittee SC13, Organization of information about construction works*.

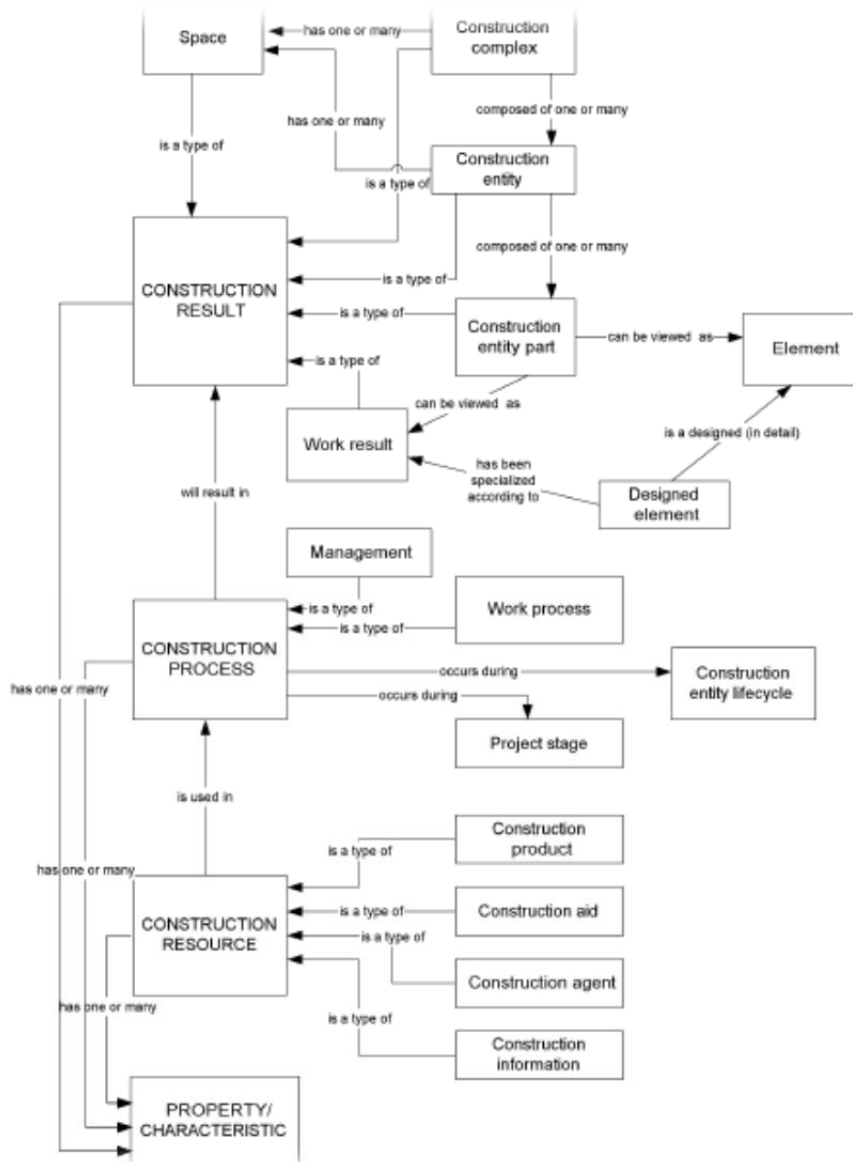


Fig. 22 – Modelo da Norma ISO 12006-2

Esta norma que sugere a seguinte proposição de base para caracterizar a indústria da construção:

A Indústria da Construção é um domínio onde se utilizam Recursos de Construção aplicáveis em Processos de Construção para obter Resultados de Construção, num processo enquadrado por Factores Técnicos.

A norma, recomenda, por outro lado, o uso de um conjunto de tabelas de classificação de auxílio à normalização, chamando a atenção para a hipótese de serem usadas com diferentes níveis de detalhe, mantendo, assim, aberta a possibilidade de adequação às necessidades potenciais de representação do conhecimento de diferentes

culturas e países, ao mesmo tempo que pretende permitir e fomentar o seu alinhamento posterior com a representação mais normalizada do conhecimento do domínio:

Provided that each country uses this framework of tables and follows the definitions given in this standard, it will be possible for standardization to develop table by table in a flexible way. For example Country A and Country B could have a common classification table of e.g. elements, but different classification tables for work results without experiencing difficulties of 'fit' at the juncture. (ISO 12006-2, 2001: 17)

O modelo proposto pela norma, que procura retratar o processo construtivo e os seus diferentes componentes, oferecendo uma visão do encadeamento dos processos e dos elementos constitutivos aplicáveis à IC, foi aceite pelos especialistas como base para a conceptualização do conhecimento do domínio da reabilitação, dada a sua origem, fundamentos teóricos, a sua grande aceitação e utilização a nível internacional, e a sua abrangência e relevância no que respeita às necessidades e propósitos do projecto H-Know.

Esteve, assim, na base do desenvolvimento do processo de conceptualização da ontologia de alto nível H-Know, tendo sido a partir do modelo e das tabelas de classificação sugeridas pela norma que o conhecimento relativo ao domínio da reabilitação foi sendo estruturado e representado, com recurso ao uso de mapas conceptuais, na fase de conceptualização.

Esta conceptualização, construída a partir da norma ISO 12006-2, foi, em seguida, enriquecida e tornada mais granular, com base na análise e reutilização da informação sobre o domínio obtida noutros recursos relevantes, recursos seleccionados e validados com a intervenção dos especialistas. Desses recursos, destacam-se:

- EPIC - Electronic Product Information Co-operation¹⁶⁶,
- Ominclass¹⁶⁷,

¹⁶⁶ <http://www.epicproducts.org> - *Electronic Product Information Cooperation - EPIC is concentrated on the definition of a common set of construction product groups including notations in order to facilitate the transfer of data between computerized national and/or distributed databases and to harmonize search patterns.*

¹⁶⁷ <http://www.omniclass.org/> - *The OmniClass Construction Classification System (known as OmniClass™ or OCCS) is a classification system for the construction industry. OmniClass is useful for many applications, from organizing library material and project information, to providing a classification structure for electronic databases. It incorporates other extant systems currently in use as the basis of many of its Tables – MasterFormat™ for work results, UniFormat for elements, and EPIC for structuring products*

- Uniclass¹⁶⁸,
- HEREIN - European Heritage Thesaurus¹⁶⁹,
- Construction Industry Knowledge Ontology – Projecto Know-Construct

Estes recursos¹⁷⁰ foram, dada a sua riqueza e abrangência, usados para a construção da conceptualização inicial de alto nível, construída a partir da norma. Algumas das fontes acima identificadas eram também multilingues, como é o caso do tesouro HEREIN, ou incluíam traduções prévias para algumas das línguas de trabalho dos especialistas da rede, como acontece com a ontologia CIK ou com parte da classificação EPIC.

Este conjunto de fontes foi disponibilizado em língua inglesa, factor que se revelou positivo também para a criação e funcionamento da rede colaborativa multilingue, para a comunicação e negociação entre os especialistas e para a gestão e representação do conhecimento no contexto da rede.

Este facto foi comprovado posteriormente, durante a entrevista semiestruturada que fizemos aos especialistas que participaram da nossa experiência¹⁷¹ que, quando questionados¹⁷² sobre se o facto de os recursos de informação e conhecimento propostos para a construção da ontologia possuírem já uma representação própria do conhecimento (ex.: hierárquica) afectava a sua representação do domínio ou, antes, a facilitava, responderam¹⁷³, na sua maioria, que facilitava, uma vez que a sua utilização: é *confortável* –

¹⁶⁸ <http://en.wikipedia.org/wiki/Uniclass> - *Uniclass is a new classification scheme for the construction industry. It is intended for organising library materials and for structuring product literature and project information. It incorporates both CAWS (Common Arrangement of Work Sections for building works) and EPIC (Electronic Product Information Co-operation), a new system for structuring product data and product literature.*

¹⁶⁹ <http://thesaurus.european-heritage.net/sdx/herein/thesaurus/introduction.xsp> - *The European Heritage Network is a permanent information system gathering governmental services in charge of heritage protection within the Council of Europe. The European heritage Network focus is on cultural heritage, particularly on architectural and on archaeological heritage.*

¹⁷⁰ O problema da adequação das fontes, resultante dos diferentes tipos e níveis de informação, bem como dos problemas no uso e na transferência do conhecimento para mapas conceptuais, foi descrito antes, no ponto 3.3.1.1.

¹⁷¹ Estes especialistas são identificados no ponto seguinte.

¹⁷² Questão número doze (ver Apêndice D).

¹⁷³ Para efeitos de análise dos resultados obtidos nas entrevistas aos especialistas, transcreveremos aqui apenas alguns segmentos das respostas obtidas. A transcrição das entrevistas pode ser consultada no Apêndice F.

poupa o especialista no processo de reflexão (JPM_{pt})¹⁷⁴, que “shortens the time to remove dilemmas” (LJU_{de})¹⁷⁵ e “helps to think about the representation and the missing information” (AA_{es})¹⁷⁶, ainda que reconheçam limitações ao seu uso, como faz (AS_{pt})¹⁷⁷, que afirma que são importantes, “mas apenas em parte, dado que há casos (domínios) que não são bem retratados nessas classificações, como é o caso da reabilitação”. Por outro lado, na sua perspectiva, “há casos em que não há uma descrição correcta do encadeamento dos processos, sendo que alguns nem chegam a ser identificados”. Esta afirmação remete-nos para os cuidados a ter no processo de adequação da informação proveniente destas fontes aos propósitos e domínios específicos a serem representados, tal como chamámos a atenção no ponto 3.3.1.1.

A análise e conjugação da informação obtida a partir dos recursos descritos conduziu a reestruturações das representações conceptuais iniciais, à medida que a pesquisa e o acesso à informação veiculada pelo *corpus* e pelos especialistas se aprofundou, num processo dinâmico de negociação e conceptualização, em que os sistemas conceptuais estavam abertos à inclusão de novos conceitos considerados relevantes e, em consequência, abertos também ao estabelecimento de novas relações e de novas perspectivas sobre os conceitos já disponíveis na representação.

Esta construção desenvolveu-se com recurso ao uso de mapas conceptuais, num ambiente colaborativo, em que os diferentes actores examinavam os problemas da representação do conhecimento do domínio a partir das suas perspectivas e visões do mundo, resultantes da formação, experiência e meio profissional de que originavam, mas também em função dos objectivos da rede colaborativa e do projecto. Da análise colaborativa e partilhada, resultou um contributo significativo para a conceptualização inicial do domínio, tendo a terminologia assumido, na nossa perspectiva, um papel relevante para o seu desenrolar.

De facto, a recolha e estruturação da terminologia do domínio da reabilitação constituiu, a nosso ver, um elemento importante para construir uma representação

¹⁷⁴ Eng. João Pedro da Silva Poças Martins, Portugal, FEUP.

¹⁷⁵ Eng. Ljubiša Urosevic, Alemanha, ATB.

¹⁷⁶ Eng. José Alberto Armijo Prieto, Espanha. TECNALIA.

¹⁷⁷ Eng. Alfredo Augusto Vieira Soeiro, Portugal, FEUP.

conceptual do conhecimento do domínio de modo mais consistente e consensual. Por outro lado, o recurso a uma abordagem interdisciplinar, que conjugou na fase de conceptualização inicial o trabalho terminológico com o trabalho de organização do conhecimento feito pelo especialista, promoveu, na nossa opinião, a uma melhor e mais eficiente elicitação e organização inicial do conhecimento do domínio.

Esta primeira fase antecedeu a aplicação do método, uma vez que foi a partir do processo de conceptualização inicial e dos mapas conceptuais dele resultantes que se iniciou o processo de especificação multilingue da ontologia e as actividades inerentes ao processo de localização, que descrevemos a seguir. Importa, no entanto, realçar que esta fase foi desenvolvida tendo em perspectiva o processo de especificação multilingue da ontologia que se seguiria, o que se reflectiu, por exemplo, na compilação do corpus textual comparável e na recolha e reutilização de recursos multilingues, como tesouros, glossários e dicionários de especialidade, entre outros.

4.2.2. Fase 2 – Definição dos participantes e da amostra

4.2.2.1. Definição dos especialistas

A selecção dos especialistas no domínio da reabilitação que participaram na experiência incidu, como seria natural, sobre os especialistas já envolvidos no projecto e na rede colaborativa H-Know. Recorremos, assim, a dois tipos de especialistas distintos: engenheiros civis de empresas da área da reabilitação e engenheiros civis ligados ao meio académico e a centros de investigação que desenvolvem a sua investigação nos domínios da construção e da reabilitação. Esta diferença no tipo de especialização dos vários especialistas, podendo representar um problema, permite, na nossa perspectiva, que os diferentes saberes em presença se complementem e confere aos processos de conceptualização e de localização uma riqueza adicional.

Os especialistas foram seleccionados também de acordo com a sua língua de trabalho, de modo a abranger os pares de línguas a serem localizados (*en-pt; en-es; en-fr, en-*

de, en-it) e a constituírem uma amostra representativa da rede. Para tal, convidámos para participar desta experiência os seguintes especialistas:

- Alfredo Augusto Vieira Soeiro, Portugal, FEUP, (AS_{pt})
- Guido Hugony, Itália, TEICOS,¹⁷⁸ (GH_{it})
- Ljubiša Urosevic, Alemanha, ATB, (LJU_{de})¹⁷⁹
- Jean-François Hornain, França, CAPEB, (JF_{fr})
- João Pedro da Silva Poças Martins, Portugal, FEUP, (JPM_{pt})
- José Alberto Armijo Prieto, Espanha, TECNALIA.¹⁸⁰ (AA_{es})

Estes especialistas foram escolhidos por possuírem um conjunto de características comuns, que considerámos relevantes para o desenvolvimento da experiência, nomeadamente:

1. São parceiros da rede colaborativa, conhecem os seus princípios de funcionamento e estão envolvidos no processo de partilha e construção do conhecimento, sendo intervenientes activos no processo de conceptualização do domínio;
2. Conhecem os objectivos do projecto H-Know, os critérios e objectivos que presidiram à construção da ontologia de domínio e os objectivos a atingir com a sua utilização na gestão do conhecimento do domínio da reabilitação;
3. Possuem dois tipos de conhecimento relevantes para a localização:
 - a. conhecimento do domínio, que se refere às características conceptuais dos objectos e estado das coisas;
 - b. conhecimento do texto, referente aos padrões de comunicação (meios verbais e/ou visuais) necessários à transmissão do conhecimento seleccionado aos receptores da localização;
4. São responsáveis por analisar a representação conceptual como um todo e aferir se esta representa o conhecimento tal qual é percebido pela comunidade profissional para a qual está a ser representado e localizado.

¹⁷⁸ Trabalhou em conjunto com a Eng. Laura Cibien, TEICOS.

¹⁷⁹ Trabalhou em conjunto com o Eng. Johannes Tuzcek, TIETJEN.

¹⁸⁰ Trabalhou em conjunto com o Eng. José Luis Izgara, LABEIN.

Para além das características descritas, os especialistas seleccionados possuem um conjunto de competências mínimas que auxiliaram na tarefa de representação do conhecimento do domínio, que provêm, para além da sua formação académica e do conjunto de saberes e conhecimentos adquiridos, do seu envolvimento no projecto e da sua exposição aos conceitos (ex.: conceito de ontologia) e às ferramentas (ex.: de construção de mapas conceptuais) inerentes a um processo de representação do conhecimento, o que os auxiliou nas actividades de eliciação, (re)conceptualização, localização e validação do conhecimento representado.

4.2.2.2. Definição da amostra

Para proceder à experimentação do método foi constituída uma amostra de mapas conceptuais a partir do caso de estudo (ver Apêndice H). Os mapas foram seleccionados tendo em vista a constituição de uma amostra abrangente. Seleccionámos doze de um total de vinte mapas, o que perfaz uma amostra de 60% do total. O número de conceitos era de cento e cinquenta e dois num total de trezentos e cinquenta e dois, correspondente a cerca de 45% do total dos conceitos.

Procurámos também que a amostra fosse representativa dos quatro módulos de alto nível definidos para a ontologia H-Know, representados na figura seguinte, uma vez que cada um representava subdomínios específicos ao domínio.

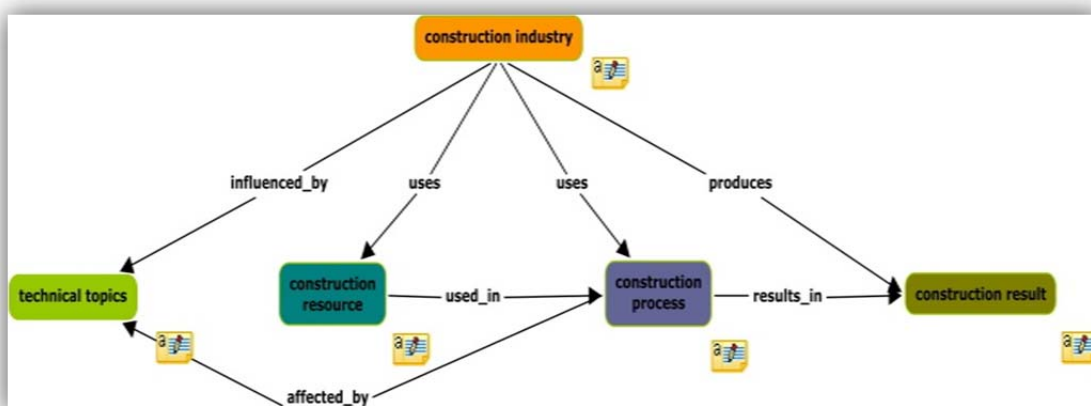


Fig. 23 – Modelo global da Ontologia H-Know

Para esse efeito, seleccionámos, de cada módulo, dois ou mais mapas, para garantir a abrangência da amostra, como se demonstra na tabela seguinte:

Amostra

Global frame of the high level ontology			
<i>Global frame of the high level ontology</i>			
Technical Topics	Construction Resource	Construction Process	Construction Result
Mapas	Mapas	Mapas	Mapas
<i>technical topic</i>	<i>construction aid</i>	<i>work process</i>	<i>construction entity</i>
<i>intervention programme</i>	<i>construction resource</i>	<i>construction process</i>	<i>construction result</i>
<i>planning system</i>	<i>material</i>	<i>project stages</i>	

Tabela 4 - Mapas conceptuais constituintes da amostra

A selecção destes mapas obedeceu a um outro critério, *i.e.*, foi elaborada de modo a incluir representações conceptuais de diferentes tipos: genéricas, partitivas e associativas. Esta selecção heterogénea foi feita de modo a podermos observar se algum dos diferentes tipos de representações conceptuais conduzia a dificuldades adicionais ao longo do processo de localização, em resultado da análise das relações conceptuais e da sua validade lógica, por um lado, e, por outro se resultavam nalgum tipo de complexidade extra quando o especialista considerava necessária uma reconceptualização e se propunha a incluir e a representar novos conceitos e as respectivas relações conceptuais.

Indicamos, em seguida, o tipo de relações conceptuais com que os especialistas se confrontaram nos diversos mapas conceptuais:

Relações conceptuais	<i>is_a</i> <i>is_a_part_of</i> <i>is_a_type_of</i> <i>is_composed_of</i> <i>is_composed_by</i> <i>influenced_by</i> <i>uses</i>
-----------------------------	--

used_in
results_in
produces
affected_by

Tabela 5: Tipo de relações conceptuais

Procurámos, finalmente, ao definir esta selecção, obter um conjunto de mapas com representações conceptuais diversificadas, algumas de menor complexidade – (ex.: mapa *construction_aid*, composto por dois conceitos e uma relação conceptual) e outros de maior complexidade (ex.: mapa *project_stages*, composto por trinta e oito conceitos e duas relações conceptuais), com o objectivo de, por um lado, verificar de que modo complexidade das diferentes representações conceptuais afectava a análise conceptual do especialista e originava sugestões de reconceptualização e, por outro, de perceber se a maior ou menor complexidade causava algum tipo de influência sobre o processo de localização.

A amostra definida¹⁸¹ revelou-se, a nosso ver, abrangente e suficiente para os propósitos estabelecidos, como explicitaremos na análise de resultados que apresentamos no ponto 4.2.5. Esta amostra esteve, como descreveremos em seguida, na origem de um conjunto de reflexões e de partilha de conhecimento que conduziram a propostas de reconceptualização de algumas das representações conceptuais do domínio, e permitiu também a observação e o acompanhamento do desenvolvimento do processo de localização neste contexto.

4.2.3. Fase 3 - Criação do espaço de conceptualização e localização

Como referimos antes, tanto o conhecimento do domínio a disponibilizar, como a representação conceptual desse conhecimento, não se poderiam restringir ao uso de uma língua em particular, tendo em conta o ambiente multilingue da rede colaborativa e a necessidade de construir uma ontologia representada em mais do que uma língua natural, pelo que a necessidade de obter uma representação do conhecimento nas diferentes línguas

¹⁸¹ Esta amostra foi definida para efeitos da experimentação do método. No entanto, todos os termos da ontologia foram localizados para as cinco línguas do projecto H-Know e explicitados formalmente no editor de ontologias Protégé (ver Anexo 4).

envolvidas no projecto se tornou manifesta no processo de construção da ontologia de alto nível H-Know, tendo a localização surgido como o processo natural a seguir para enriquecer esta ontologia com a componente multilingue.

O processo de conceptualização foi, como aclarámos antes, desenvolvido inicialmente em língua inglesa, num ambiente de especificação semiformal da conceptualização, com recurso ao uso de mapas de conceptuais, de modo a permitir e a potenciar a participação de todos os actores da rede colaborativa. Nesse sentido, e tendo em conta a necessidade de permitir que a localização se realizasse num espaço familiar e acessível aos especialistas, desenvolvemos um cenário de trabalho colaborativo de apoio à conceptualização e localização, tal como proposto no ponto 3.7.1.

Assim, para dotar o ambiente de trabalho dos elementos necessários à participação dos especialistas na análise das representações conceptuais e na sua localização, e para facilitar a sua acção, recorremos aos mapas conceptuais¹⁸², desenvolvidos em CmapTools, e adicionámos um conjunto de elementos específicos para apoiar o processo de localização, que permitissem integrar os princípios e métodos terminológicos no contexto da situação comunicativa específica e da aplicação particular, o que é, em seguida, exemplificado:

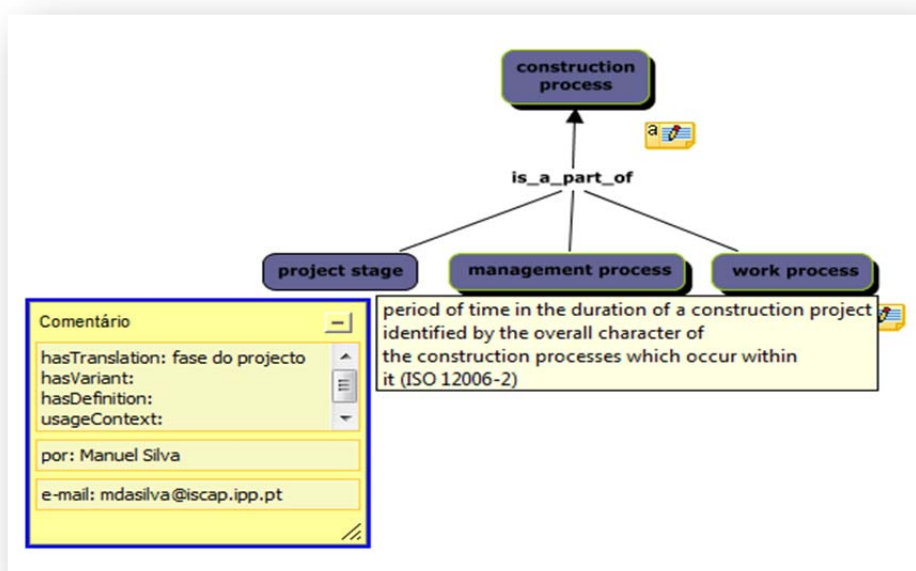


Fig. 24 - Ex.: Espaço de conceptualização e localização - Mapa *construction process*

¹⁸² A análise do contributo resultante da utilização de mapas conceptuais para a representação do conhecimento foi desenvolvida no ponto 3.6.

Criámos, para tal, a possibilidade de o especialista – ou do terminólogo - associar a cada conceito do mapa uma ou mais definições em língua natural. Procurámos, por outro lado, através do uso do sistema de comentários, permitir a inserção e, posteriormente, o acesso dos especialistas aos equivalentes nas diferentes línguas de trabalho. Foi criado, para este efeito, um espaço que permitia a inclusão do seguinte conjunto de elementos: 1) *equivalente*, 2) *uma ou mais variantes*, 3) *definições* e 4) *contexto de uso*. Estes elementos, uma vez completos e validados, poderiam ser utilizados no processo de especificação multilingue da ontologia em termos formais.

Como descrevemos antes, previmos também a possibilidade do processo de localização e de enriquecimento dos mapas conceptuais com informação multilingue poder decorrer de dois modos:

1. a partir da sugestão/inclusão prévia de equivalentes feita pelo terminólogo, obtidos com base nos recursos de especialidade multilingues previamente identificados e validados (dicionários, terminologias, classificações, tesouros e ontologias), sendo estes equivalentes posteriormente revistos e validados pelo especialista,
2. a partir da sugestão/inclusão feita pelo próprio especialista, com base na sua experiência e no conhecimento que possui da sua língua de especialidade.

Assim, e de modo a agilizar o processo de localização, incluímos, enquanto participantes deste processo, em cada mapa as definições dos conceitos em língua inglesa. A inclusão das definições foi feita com o objectivo de permitir o acesso dos especialistas a uma definição previamente validada de cada conceito, processo concluído durante a conceptualização inicial, de modo a terem essa definição presente no desenvolvimento do processo de localização e na escolha dos equivalentes. Incluímos também, para um grande número de termos dos diversos mapas conceptuais, uma primeira proposta de equivalente. Quer as definições em língua natural quer os equivalentes propostos foram obtidos a partir dos recursos do domínio anteriormente validados, nomeadamente da norma ISO 12006-2, das classificações internacionais Uniclass e Omniclass, do tesouro HEREIN e da ontologia CIK.

Não foram sugeridos equivalentes para a totalidade dos termos¹⁸³, de modo a colocar os especialistas perante dois tipos de cenários, nomeadamente: aquele em que já é sugerido um equivalente, que os especialistas podem ou não aceitar/rever; aquele em que a pesquisa e selecção do equivalente dependem inteiramente da acção dos especialistas. Com esta metodologia pretendeu-se observar como agiam os especialistas em cada situação, quais os problemas suscitados por cada termo e quais as estratégias usadas para a sua localização.

4.2.3.1. Definição dos recursos e aplicações de apoio à localização

Para implementar o método e facilitar a participação do especialista no processo de localização procedemos à selecção de um conjunto de recursos e de aplicações web com o objectivo de sustentar a pesquisa e selecção dos equivalentes dos termos. Para a escolha daqueles recursos e aplicações tivemos em conta os critérios enunciados no ponto 3.7.2, como: o conteúdo a localizar; o contexto em que o processo decorreria; as competências previsíveis dos especialistas e os objectivos a atingir com o processo de localização.

Assim, após uma análise crítica ao funcionamento das aplicações da *Web* de tradução e localização, que considerou quer as necessidades de localização de um domínio de especialidade quer a acessibilidade das aplicações e a sua facilidade de uso por parte dos especialistas, identificámos e seleccionámos um conjunto de aplicações de tradução automática e de bases de dados terminológicas e lexicais bilingues e multilingues disponíveis em linha.

O processo de localização não se desenrolou, no entanto, com recurso a apenas aquelas aplicações. De facto, este processo é, necessariamente, influenciado pelo conhecimento pré-existente relativo ao domínio, que é acrescentado ao processo pelo especialista ou que resulta da análise e reutilização de recursos de conhecimento do domínio devidamente validados, como já referimos. Com efeito, ainda que possamos socorrer-nos de aplicações de tradução automática e tesouros, é importante, em domínios específicos, encontrar outros meios apropriados de apoio à tradução e à localização

¹⁸³ No caso da língua italiana, o número de equivalentes sugeridos para análise e validação foi bastante menor, dada a inexistência de traduções prévias nos recursos identificados e os poucos recursos multilingues de especialidade encontrados para o domínio em análise.

especializadas, como glossários e dicionários multilingues de especialidade, que contribuam também para identificar e seleccionar os equivalentes mais apropriados para um determinado termo.

As ferramentas e os recursos disponibilizados aos especialistas foram diferentes de língua para língua, em número e em tipo, uma vez que nem todas as línguas possuem o mesmo tipo de recursos, nem estes o mesmo grau de desenvolvimento e de actualização. Apresentamos, abaixo, um exemplo das ferramentas seleccionadas para a língua portuguesa, em que se distinguem as de uso mais genérico e universal, como as aplicações de tradução automática, e outras mais específicas ao domínio e à língua de trabalho, como é o caso dos glossários e dicionários de especialidade e sítios *Web* de referência sobre temáticas do domínio, como o Patorreb¹⁸⁴, bem como sítios da IC já localizados.

FERRAMENTAS DE APOIO À LOCALIZAÇÃO

Aplicações *Web* de tradução automática

Google Translator	http://translate.google.pt/
Bing Translator	http://www.microsofttranslator.com/
Babel Fish	http://babelfish.yahoo.com/
Linguee	http://www.linguee.pt/

Bases de dados terminológicas:

IATE (BDT Comunidade Europeia)	http://iate.europa.eu
--------------------------------	---

Bases de dados lexicais:

Wordnet	http://wordnet.princeton.edu/
Wordnet Português	http://www.clul.ul.pt/wn/

¹⁸⁴ <http://www.patorreb.com/pt/> Ver exemplo de ficha no anexo 3.

Motores de pesquisa de traduções:

LazyTerm	http://terminotrad.info/RogerMcKeon/LazyTerm/TerminoParesse-Rev.3e.html
Global Glossary	http://en.globalglossary.org/Default.aspx

OUTROS RECURSOS:

Glossários e dicionários do domínio - PT

Dicionário da Construção Civil	http://www.ecivilnet.com/dicionario/
Construção Civil -Glossário de termos técnicos	www.forma-te.com/
Dicionário abc da construção	http://luispatricio.planetaclix.pt/abconstrucao.htm

Recursos específicos do domínio

Patorreb	http://www.patorreb.com/pt/
----------	---

A estas ferramentas e recursos foram sendo acrescentados outros ao longo do desenvolvimento do processo de localização, por proposta dos especialistas, como:

Páginas web localizadas

(páginas localizadas em diferentes línguas na área da comercialização de produtos de construção)

SIKA	http://prt.sika.com/pt/solutions_products/02.html
Grupopuma	http://www.grupopuma.com/pt/

Esta selecção procurou ter em conta as características dos participantes, que, como podemos perceber pela análise das entrevistas¹⁸⁵, traduzem quase todos, ainda que de modo não muito frequente, usando para esse fim sobretudo dicionários gerais e específicos disponíveis em linha e motores de tradução automática, destacando-se o uso do motor de tradução automática do Google, uso a que atribuem um grau de dificuldade baixo ou médio e cujas traduções obtidas descrevem como tendo uma qualidade média. Nenhum especialista recorreu, no âmbito deste estudo, a bases de dados terminológicas ou lexicais¹⁸⁶, sendo a sua existência desconhecida de todos até ao início deste processo.

Finalmente, tendo sido questionados¹⁸⁷ sobre se as ferramentas de localização que utilizaram produzem resultados com qualidade suficiente para auxiliar um especialista sem formação em tradução a participar de modo eficiente no processo de localização de uma ontologia, a resposta foi positiva por parte de todos os inquiridos, tendo (AA_{es}) salientado que a *“technology is evolving and engines are more sophisticated and improving. But results depend on the features you want to use. You get better results by using specific domain tools”*. Todos destacaram, finalmente, o auxílio inicial que o uso daquelas ferramentas representou para o processo.

Como pudemos constatar, no entanto, o recurso a aplicações como tradutores automáticos, apesar do seu constante avanço e melhoria, continua a oferecer resultados de baixa ou média qualidade, sobretudo no caso de línguas com recursos ainda não tão robustos, como se verificou no caso da língua italiana, podendo o seu uso exclusivo tornar-se uma limitação para o desenvolvimento do método que propomos. Torna-se, assim, importante, alertar os especialistas para a diversidade e riqueza dos recursos existentes e disponíveis para os auxiliar, sobretudo os específicos ao domínio de especialidade.

¹⁸⁵ Questões dois a cinco (ver Apêndice F).

¹⁸⁶ O seu uso foi no entanto proposto e demonstrado pelo terminólogo em diferentes momentos.

¹⁸⁷ Questão número sete (ver Apêndice F).

4.2.4. Fase 4 – Experimentação do método

4.2.4.1. Desenvolvimento de um guião para o especialista

Definidos os recursos, o ambiente de trabalho, os especialistas e a amostra, foi elaborado um guião para auxiliar e documentar o especialista no desenvolvimento da análise conceptual e da localização (ver Apêndices B e C). Este guião explicita os objectivos da participação do especialista e das tarefas a desenvolver e foi criado com a intenção de o dotar de um instrumento a que pudesse recorrer de modo sistemático em caso de dúvidas ou para registar as suas reflexões.

Para esse fim, o guião disponibilizava informações sobre o tipo de acções a empreender e os recursos disponíveis, quer para a análise da conceptualização e eventual reconceptualização, quer para a sua localização, tendo-se revelado útil também como complemento à acção do terminólogo. No guião, de modo a clarificar o especialista sobre o tipo de interacção a desenvolver com o conjunto de recursos a que acederia, identificámos e descrevemos:

- o funcionamento do espaço de conceptualização e localização;
- a ferramenta de apoio à representação conceptual;
- as ferramentas de apoio à localização;
- outros recursos de apoio ao desenvolvimento da localização, nomeadamente os recursos do domínio já validados;
- os restantes recursos de apoio¹⁸⁸.

O guião foi ainda elaborado com o intuito de dotar o especialista de um espaço onde colocar as suas reflexões sobre a análise feita a cada mapa, os problemas encontrados e as suas dúvidas ou sugestões específicas para a resolução de cada um dos problemas, quer de localização quer de (re)conceptualização, servindo igualmente para a inserção de comentários considerados relevantes pelo especialista para o bom desenvolvimento das tarefas em curso ou de outras que pudessem ultrapassar o âmbito do processo, mas cuja

¹⁸⁸ Todos os recursos existentes em formato electrónico foram, também, enviados ao especialista. Entre estes, incluíam-se, por exemplo, as classificações e os dicionários e glossários de especialidade.

reflexão o especialista achasse conveniente partilhar. Nesse sentido, foi disponibilizada, na parte final do guião, uma lista dos mapas a conceptualizar e um espaço para recolha de observações, sendo esse espaço antecedido de uma descrição sobre o tipo de comentários que o especialista poderia sugerir.

O guião foi, então, enviado aos especialistas juntamente com a amostra, composta pelas representações iniciais em língua inglesa e pelos recursos relevantes para a estruturação do conhecimento do domínio, tendo cada especialista recebido recursos específicos ao par de línguas sobre o qual se debruçou.

4.2.4.1.1. Acompanhamento do especialista na aplicação do método

O envio do guião foi precedido, como seria natural, de um convite a cada especialista e de uma primeira reunião. De facto, esta fase pressupõe que todos os especialistas sejam contactados pelo investigador, de modo a esclarecer os objectivos da experiência e os passos a seguir; a identificar e descrever o ambiente de conceptualização e localização e a explicitar o conjunto de recursos disponíveis para a realização da experiência, bem como o papel de cada um.

À primeira reunião sucederam-se um conjunto de reuniões periódicas ao longo do desenvolvimento de todo o processo, que nos permitiram acompanhar de forma interactiva e contínua o especialista e observar o desenrolar do processo. Nestas reuniões, foram clarificadas dúvidas sobre a realização da experiência e fornecidas orientações ao especialista em função das suas necessidades. Estas reuniões tiveram lugar em função da disponibilidade e/ou da necessidade dos especialistas, tendo sido realizadas presencialmente, via Skype ou por telefone¹⁸⁹, e serviram o intuito de:

1. clarificar dúvidas sobre o uso das aplicações de conceptualização e de localização;
2. clarificar dúvidas sobre como proceder em relação às alterações que o especialista pretendesse propor e que considerasse relevantes em termos de cada mapa conceptual;

¹⁸⁹ Sempre que solicitado, foram também esclarecidas dúvidas por correio electrónico.

3. detectar as dificuldades e problemas sentidos pelos especialistas na (re)conceptualização e localização dos termos do domínio e auxiliar na sua resolução;
4. analisar as estratégias e ferramentas utilizadas na resolução das dificuldades com que os especialistas se depararam e, quando necessário, sugerir o uso de alternativas;
5. observar e analisar o resultado da participação dos especialistas no processo de (re)conceptualização e localização das representações conceptuais;
6. obter informações sobre a sua percepção quanto ao processo em curso e à metodologia seguida.

As reuniões serviram, finalmente, para tornar claro o papel de representação semiformal do conhecimento desempenhado pelos mapas conceptuais, tendo sido estabelecida e exemplificada a distinção entre este tipo de representação e uma representação formal, com recurso a exemplos baseados em ontologias implementadas no editor de ontologias Protégé, onde a ontologia H-Know seria implementada formalmente.

Estabelecer esta distinção permitiu, na nossa perspectiva, tornar mais claras as tarefas a desenvolver nas diferentes etapas desta experiência e concretizar, quer as estratégias a usar na análise conceptual e possível reconceptualização quer a própria natureza dos resultados a obter. Permitiu, por outro lado, uma maior coerência nos dados que recolhemos para análise e que analisaremos em seguida.

4.2.5. Fase 5 – Análise de resultados

Após a conclusão do processo descrito no ponto anterior e a recepção dos mapas e comentários dos especialistas, procedemos à análise dos resultados, a partir da observação das decisões de localização, realizadas a partir das representações conceptuais iniciais, e das propostas de (re)conceptualização, construídas com base no resultado do processo de localização para as diferentes línguas de chegada. Esta análise de resultados foi levada a cabo com base num estudo exploratório-descritivo (Fortin, 2003: 240-242) em que se conjugaram diferentes técnicas para a colheita dos dados, nomeadamente:

1. a observação directa, já descrita, resultante da nossa participação enquanto investigador participante no projecto de investigação;
2. a realização de uma análise, partindo de um ponto de vista terminológico, aos comentários, às sugestões e às escolhas propostas pelos especialistas durante o processo de análise das representações conceptuais e o processo de localização dos termos representados em cada mapa conceptual;
3. a realização de uma entrevista semiestruturada aos especialistas (face a face, via Skype e por telefone).

4.2.5.1. Análise dos resultados da experimentação do método

É nossa perspectiva que a escolha dos conceitos que constituem uma representação do conhecimento de um domínio e a sua localização estão interrelacionadas e derivam da própria natureza das ontologias, sendo que ambas dependem da abordagem desenhada para o desenvolvimento de cada ontologia. Como explicitámos antes, o objectivo mais genérico do desenvolvimento de processos de especificação multilingue de ontologias, pelo recurso à localização, é o de permitir a interoperabilidade semântica interlinguística em sistemas de informação que contêm, normalmente, um grande número de recursos de conhecimento heterogéneos. Localizar uma ontologia é, por isso, um processo com problemas próprios, já descritos. Quando o ponto de partida é um mapa conceptual podem surgir dificuldades adicionais, uma vez que o especialista se confronta, em paralelo, com a representação do conhecimento especificada em cada mapa conceptual e com a localização dos termos neles representados.

As tarefas principais que acompanharam o processo de localização a partir de mapas conceptuais consistiram na identificação e validação dos conceitos e na localização dos termos que os designam; na análise das relações que estabelecem os conceitos no que se refere à sua validade lógica (e à validade lógica da representação conceptual no seu todo) para uma língua e um contexto profissional particular; e pela análise da adequação do conhecimento representado ao processo de pesquisa e gestão de informação, que esteve na génese da construção da ontologia e no qual esta seria utilizada.

Foi, assim, solicitado aos especialistas que procedessem à análise das representações conceptuais iniciais disponíveis em língua inglesa e localizassem os termos ou revessem/validassem as propostas de equivalente já disponibilizadas nos mapas conceptuais, tal como demonstra o exemplo seguinte.

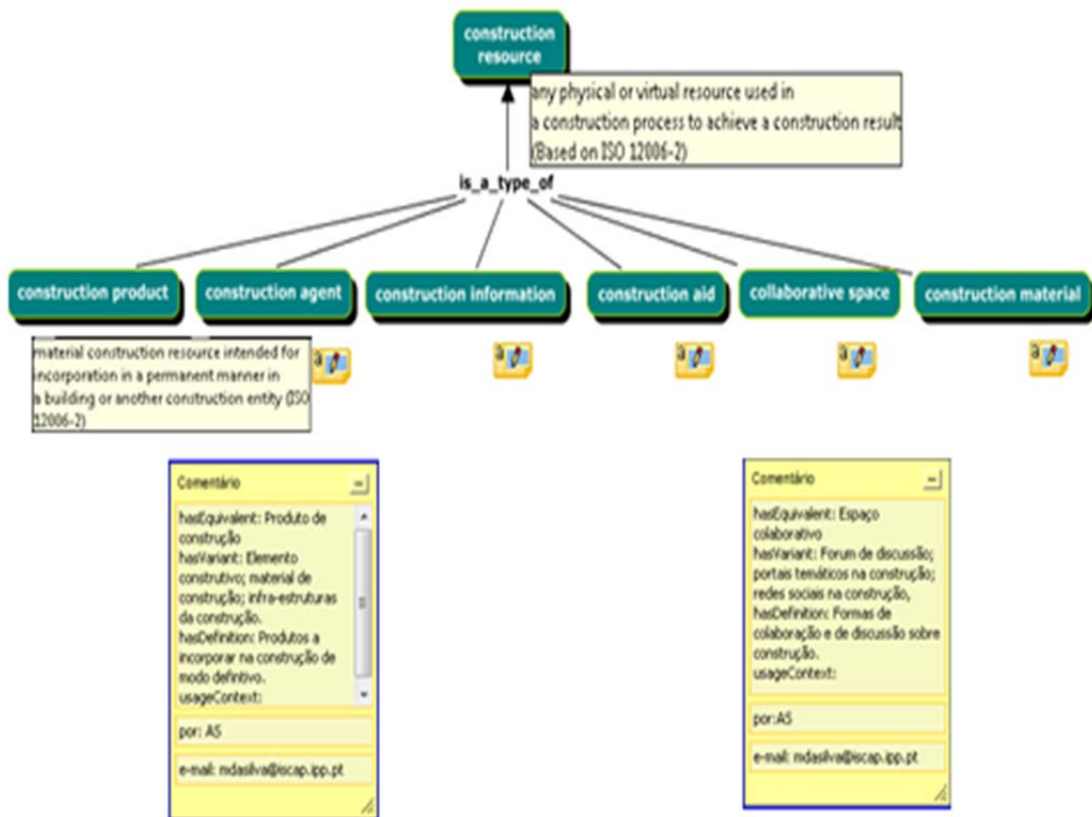


Fig. 25 – Espaço de conceptualização e localização: exemplo de uso

Para este efeito, foram propostas diferentes estratégias e ferramentas de apoio à resolução dos problemas de localização e de representação multilingue da conceptualização, anteriormente descritas. Os especialistas, durante o desenvolvimento dos dois processos, o que ocorreu de modo consecutivo no espaço de conceptualização e localização, foram colocados perante dois tipos de situações, nomeadamente (1) as que conduziram a alterações na representação conceptual e (2) as que não conduziram a qualquer alteração.

4.2.5.1.1. Desenvolvimento do processo de localização

O processo de localização partiu, como descrevemos, de uma conceptualização inicial desenvolvida e representada em língua inglesa sob a forma de mapas conceptuais. O uso da língua inglesa como língua pivô tende a facilitar, até certo ponto, o processo de comunicação e, neste caso, de localização, quando estamos perante uma comunidade multilingue, mas coloca também várias dificuldades, como pudemos observar, muitas das quais transversais aos restantes processos de localização e tradução.

O processo de localização partiu também da reutilização de traduções já existentes ou propostas no âmbito de outros recursos multilingues validados como relevantes para o domínio e para o projecto, de que são exemplo os módulos adaptados e reutilizados a partir da ontologia CIK (Construction Industry Knowledge Ontology) e do European Heritage Taurus. Os termos e respectivos equivalentes disponibilizados nestes recursos foram propostos para a ontologia e integrados nos mapas conceptuais após uma primeira análise dos propósitos e da abordagem seguida para a concepção de cada uma das fontes, elaborada de modo perceber qual o contexto de construção do recurso, a sua autonomia ou dependência em relação a este, e qual a metodologia que conduziu à escolha dos equivalentes (no caso da existência de documentação sobre esse processo).

Esta selecção inicial de equivalentes foi desenvolvida por nós, enquanto investigador participante do projecto, de modo a agilizar o desenvolvimento do processo de localização. Àqueles equivalentes acrescentámos algumas propostas de traduções obtidas a partir de glossários e dicionários multilingues de especialidade¹⁹⁰, propostas que, como pudemos verificar, foram alvo de um processo de revisão cuidadosa, de que daremos exemplo em seguida. Como é natural, para um conjunto de termos não encontramos proposta de equivalente, tendo a selecção destes ficado a cargo de cada especialista.

¹⁹⁰ Ex.: Dictionary of Building and Civil Engineering (en-fr): em <http://www.casaplusconstruction.com/mo/dic.pdf>; Technical glossary for structural engineering (en-de-fr-it-pt): em http://issuu.com/onomatranslations/docs/glossary_structural_engineering_multi_languages

Assim, no âmbito do processo de localização que se seguiu, os equivalentes propostos foram alvo de uma nova revisão¹⁹¹ com vista à actualização e à validação da terminologia à luz do conhecimento específico do domínio dos especialistas. No entanto, estes nem sempre conseguiram resolver os problemas com que se depararam, apesar do seu conhecimento do domínio.

A localização de uma representação do conhecimento coloca tanto especialistas, como terminólogos, perante dificuldades que podemos descrever como transversais a todos os processos de localização, tal como a necessidade de determinar:

1. se a sua língua de trabalho (= língua de chegada) dispõe de uma unidade terminológica lexicalizada para expressar cada um dos conceitos presentes na representação de partida, ou se é necessário recorrer a um termo com um significado próximo para representar essa realidade. Essa dificuldade está expressa no comentário que se segue, em que LJU_{de}, a propósito do mapa *construction_entity* e do termo *construction element*, cujo equivalente proposto fora *Baueinheit* afirma:

“Baueinheit” is something as “construction component” not really built object, corresponds to the ISO 12006 “construction entity part” - valid also for construction_result.

Este caso espelha uma situação comum a outros termos (e outras línguas), em que a designação do conceito reflecte apenas um subconjunto de características que não assumem, sistematicamente, o mesmo valor de uma língua para a outra.

Um outro exemplo é dado por AS_{pt} no mapa *work_process* em relação ao conceito *prefabricated buildings and structures installation*, sobre o qual afirma: *“Não encontro tradução adequada. Não percebo exactamente o significado da expressão”*. Apesar do termo constar de uma classificação internacional (OmniClass), seria, na opinião do especialista, mais adequado a existência de dois termos: *“prefabricated buildings installation”* e

¹⁹¹ Apresentamos aqui um dos exemplos de uma proposta de revisão, feita por LJU_{de} para o mapa *Intervention_programme*:

- *Intervention programme should be “restoration project plan” and German “Sanierungsplan”,*
- *“Regeneration schemes” should be “work plan” or “work plan documents” and German “Bauarbeiten Plan” or “Bauarbeiten Plan Dokumente”,*
- *Both “programme” I would replace with “plan” and correspondingly translate it with “Schutzplan” and “Restorationplan”.*

“*prefabricated structures installation*”, o que facilitaria também o processo de escolha de um equivalente¹⁹².

Este exemplo¹⁹³, entre outros que pudemos observar, demonstra que os termos usados para designar um determinado conceito num sistema conceptual (ex.: termos oriundos de classificações internacionais) podem não ser completamente entendidos ou enquadráveis à luz de uma determinada experiência ou contexto profissional e esta incompreensão levar o especialista a não encontrar equivalentes exactos na terminologia do domínio de uma língua de chegada, ainda que esta possa eventualmente já existir.

Esta diferença ao nível das características de alguns conceitos, demonstra que, apesar de estarmos perante uma área da engenharia muitas vezes vista como um domínio internacionalizado¹⁹⁴, para a qual se podem encontrar traduções para a generalidade dos termos que designam os “*concepts in the ontology, since the same conceptualization is shared among the languages represented in the ontology*” (Montiel-Ponsoda (2011: 05), este nem sempre é, de facto, o caso, quando nos referimos à engenharia civil e à indústria da construção, onde existem, como descrevemos antes, influências culturais no modo de concretizar os processos construtivos, pelo que as características de alguns conceitos diferem, de facto, de cultura para cultura e de língua para língua, sendo necessário resolver estes problemas de modo a dotar a ontologia dos equivalentes adequados.

De facto, enquanto a relação de equivalência entre conceitos genéricos como *construction process* (en) = *processus de construction* (fr) = *proceso constructivo* (es) = *processo de construção* (pt) = *Bauprozess* (de) = *processo edilizio* (it) é, até certo ponto, relativamente previsível, uma vez que os conceitos se recobrem de uma língua para a outra, os conceitos com características mais específicas, como *prefabricated buildings and structures installation*, tornam-se, por vezes, mais difíceis de analisar do ponto de vista

¹⁹² O especialista optou, posteriormente, pelo equivalente “*Instalação de pré-fabricados*”. Nesta situação, e em situações análogas, recorreremos ao corpus português como forma de tentar verificar a existência do equivalente proposto, de modo a auxiliar o especialista na sua escolha e nas suas decisões. (Para este caso, não foi encontrada qualquer ocorrência.)

¹⁹³ Este exemplo coloca-nos perante a questão: se o especialista não consegue resolver a questão da equivalência, poderá o recurso directo e exclusivo a aplicações (semi)automáticas obter, de facto, melhores resultados? Ou contribui antes para aumentar a ambiguidade e o ruído neste processo?

¹⁹⁴ Domínios que Montiel-Ponsoda (2011: 05) descreve como “*technical or specialized domains of knowledge such as engineering or medicine that have standards for processes and descriptions, and whose categorizations usually reflect the common view of different cultures*”

conceptual, uma vez que acedemos a diferentes perspectivas sobre o mesmo objecto, a que acresce o facto de as designações dos conceitos não se reflectirem directamente de uma língua para a outra, uma vez que manifestam, por vezes, marcas específicas de cada língua.

2. se a equivalência entre termos é possível em casos que são marcados em termos culturais, tal como descreve o especialista LJU_{de} no comentário seguinte:

(...) « construction agent » is a human being in ISO and in German can be both “mittel” – means, and somebody who is active on the construction site “Site agent = Bauleiter”.

Neste caso, o termo atribuído ao conceito em língua alemã não permite estabelecer uma correspondência exacta, dado que pode ser utilizado para referir mais do que uma realidade, o que não acontece nas outras línguas. Como percebemos pelo exemplo, a aparente equivalência directa (um para um) entre termos de duas ou mais línguas nem sempre é viável, dado que o termo que designa o conceito representa características com valores que diferem entre línguas, o que causa problemas na localização do sistema conceptual podendo conduzir a dificuldades em fazer coincidir os elementos de dois sistemas conceptuais de línguas diferentes, em resultado da não uniformidade do significado nas diferentes línguas. Estas dificuldades, presentes também na localização de outros termos quer para a língua alemã quer para as restantes, demonstram parte da complexidade inerente ao processo de localização.

A falta de um termo que permita a localização pode, por outro lado, causar problemas na compreensão da representação conceptual, levando a que o especialista, por exemplo, possa questionar a necessidade de inclusão do conceito na ontologia.

Observamos, a partir do acompanhamento do processo de localização e da recolha de exemplos como os acima descritos, que a análise e eventual reestruturação da representação conceptual e a explicitação das dúvidas, reforçada pela necessidade de desenvolver em simultâneo a localização, potenciam a consciencialização do especialista, ao confronta-lo com a necessidade de explicitar as suas questões e incertezas, a necessidade de consultar as fontes de conhecimento comuns e validadas (ex.: classificações e normas) e a necessidade de se posicionar face a estas – do seu ponto de vista, do ponto de vista do seu

contexto profissional e, finalmente, do ponto de vista da sua língua, como se pode observar no comentário seguinte, de LJU_{de} sobre o mapa *construction_aid*:

“Construction aid” (according to the sources I was able to find) means “financial support” for building objects and in that respect I cannot see direct relationship in the concepts. According to the ISO 12006 it is a kind of support from machines and constructions but again I was not able to find appropriate German translation.¹⁹⁵

Como pudemos observar a partir do exemplo, os conceitos expressos numa ontologia podem não ter uma correspondência directa com as formas lexicais específicas de uma determinada língua natural, o que demonstra a dificuldade, existente em alguns casos, em estabelecer uma equivalência directa entre o nível linguístico e o nível conceptual.

Para tal tornou-se importante a existência, no espaço de conceptualização e de localização que concebemos, da possibilidade do especialista introduzir variantes à sua proposta, de modo a poder transmitir todas as possibilidades que identificou/considerou para cada termo, como aconteceu nos exemplos seguintes:

JPM_{pt} – Mapa *construction_resource*:

investigation_and_survey

hasTranslation: investigação e análise

hasVariant: levantamentos

AS_{pt} – Mapa *intervention_programme*:

regeneration_schemes

hasTranslation: Programas de regeneração

hasVariant: Programa de reconversão ou adaptação de construção.

hasDefinition: programas destinados a modificar a construção.

AS_{pt} – Mapa *construction_entity*:

civil_engineering_product

hasTranslation: Obra de engenharia civil

hasVariant: Obra pública; Infra-estrutura

hasDefinition: Obra de utilização pública

¹⁹⁵ O equivalente proposto foi, após análise de diferentes opções, “*Konstruktionshilfe*”, com base em <http://www.dict.cc/?s=subject:constr.>

AS_{pt} – Mapa *construction_resource*:

construction_product

hasTranslation: Produto de construção

hasVariant: Elemento construtivo; material de construção; infra-estruturas da construção.

hasDefinition: Produtos a incorporar na construção de modo definitivo.

Esta possibilidade demonstra também a atitude crítica, atenta e activa dos especialistas, que percebem a necessidade de propor soluções, ainda que estas não os satisfaçam totalmente, de modo a dotar a ontologia de uma componente multilingue. Noutros casos a possibilidade de introduzir variantes foi utilizada, como pudemos observar nos exemplos que se seguem, para propor novos conceitos, mais específicos, que poderiam vir a fazer parte da ontologia e a torná-la mais granular, ou vir a ser usados numa ontologia de nível local (tal como definido no âmbito do projecto H-Know).

AS_{pt}¹⁹⁶ – Mapa *construction_entity*:

recreational_and_entertainment_facility

hasTranslation: Equipamento de lazer

*hasVariant*¹⁹⁷: Estádio; Cinema; Centro desportivo; Sala de jogos.

hasDefinition: Construção dedicada a actividades desportivas ou de entretenimento.

AS_{pt} – Mapa *construction_resource*:

construction_agent

hasTranslation: Agente de construção

hasVariant: Interveniente na construção; Empreiteiro; Dono de Obra; Fiscal; Projectista.

hasDefinition: Entidade pública ou privada envolvida no processo da construção.

¹⁹⁶ Destacamos aqui os exemplos de AS_{pt}, por serem os mais ricos e mais completos para os efeitos desta análise.

¹⁹⁷ Como resulta da análise dos exemplos, o elemento *hasvariant* foi utilizado para indicar não só potenciais sinónimos, mas também hipónimos.

AS_{pt} – Mapa *work_process*:

groundwork

hasTranslation: Fundações e movimento de terras

hasVariant: Sapatas; caboucos, aterros; escavações.

*hasDefinition: Trabalho realizado no terreno ou no solo que envolve escavações ou movimento de terras*¹⁹⁸

No que se refere às definições, para além das disponibilizadas na representação conceptual inicial, foi feito apenas um pequeno número de propostas na língua de chegada, sendo que as características das novas definições as aproximavam do que designaremos por descrições funcionais, tendo, quase todas, sido criadas pelos próprios especialistas, com propósitos descritivos¹⁹⁹ ou de contextualização, como é o caso das definições propostas nos três exemplos acima transcritos.

Ao longo do desenvolvimento do processo de localização, a informação introduzida pelos especialistas contribuiu também para aprofundar a análise terminológica e conceptual do domínio e revelou-se um auxílio importante para o terminólogo, ao permitir uma compreensão mais profunda sobre o domínio, obtida a partir da análise da nova informação e das novas perspectivas expressas pelos especialistas, bem como da análise das fontes e da documentação adicionais que estes identificaram e propuseram como relevantes, mesmo quando estas descreviam uma realidade marcadamente nacional. Esta documentação adicional contribuiu, assim, para uma monitorização mais completa das fontes relevantes para o domínio, para a sua melhor compreensão, e para uma representação mais granular e efectiva do domínio.

Como percebemos pelos exemplos e comentários transcritos, a presença do especialista permite, em parte, minorar alguns dos problemas inerentes à celeridade e

¹⁹⁸ Dos sete exemplos destacados, quatro continham uma proposta de partida para a localização seleccionada pelo terminólogo, e três não continham qualquer proposta. Das propostas iniciais apenas uma sofreu modificações. Todos os restantes elementos (variantes e definições) são da autoria do especialista.

¹⁹⁹ Roche (2012: 27) identifica dois tipos de definição “*the more exact one, which retains the name definition, and the other, less exact, which is called a description. The more exact definition explains the nature of a thing by its essential attributes, of which the common one is called the genus, and the proper one the difference. The less exact definition, called a description, provides some knowledge of a thing in terms of the accidents that are proper to it and determine it enough to give us an idea distinguishing it from other things*”.

efectividade da representação do conhecimento especializado no contexto de um processo de localização. De facto, a sua presença contribui para a resolução (1) de problemas conceptuais, dado que conhece o domínio, o que contribui para ultrapassar mais facilmente os problemas resultantes da ambiguidade inerente a cada língua; (2) de problemas linguísticos, uma vez que tem competência na língua de especialidade e reconhece a quase totalidade dos termos a localizar, simplificando assim, em grande parte dos casos, a resolução do problema da possível ausência de equivalentes; (3) de problemas pragmáticos, relativos ao uso do termo (alcance geográfico, nível de normalização, grau de aceitação, entre outros), que consegue mais facilmente definir e prever.

4.2.5.1.2. Desenvolvimento do processo de (re)conceptualização

Ao longo do desenvolvimento do processo de (re)análise conceptual, pudemos observar que os especialistas se socorreram, para a levar a cabo, sobretudo de três processos: a *adição*, a *modificação* e a *eliminação* de conceitos. Estes processos resultaram num conjunto de contributos e propostas de reconceptualização que exemplificamos e analisamos em seguida.

Destacaremos, para esta análise, uma amostra de um número necessariamente limitado de exemplos, recolhida a partir dos contributos e comentários dos especialistas, dado que seria impossível representar ou transcrever para este trabalho todos os dados recebidos.

1. **Adição de conceitos** à representação conceptual. A adição assume diferentes formas:
 - a. **adição de novos conceitos e/ou relações conceptuais**, relevantes do ponto de vista do especialista e da sua realidade nacional ou local, tal como proposto nos mapas abaixo:

Exemplo 1:²⁰⁰

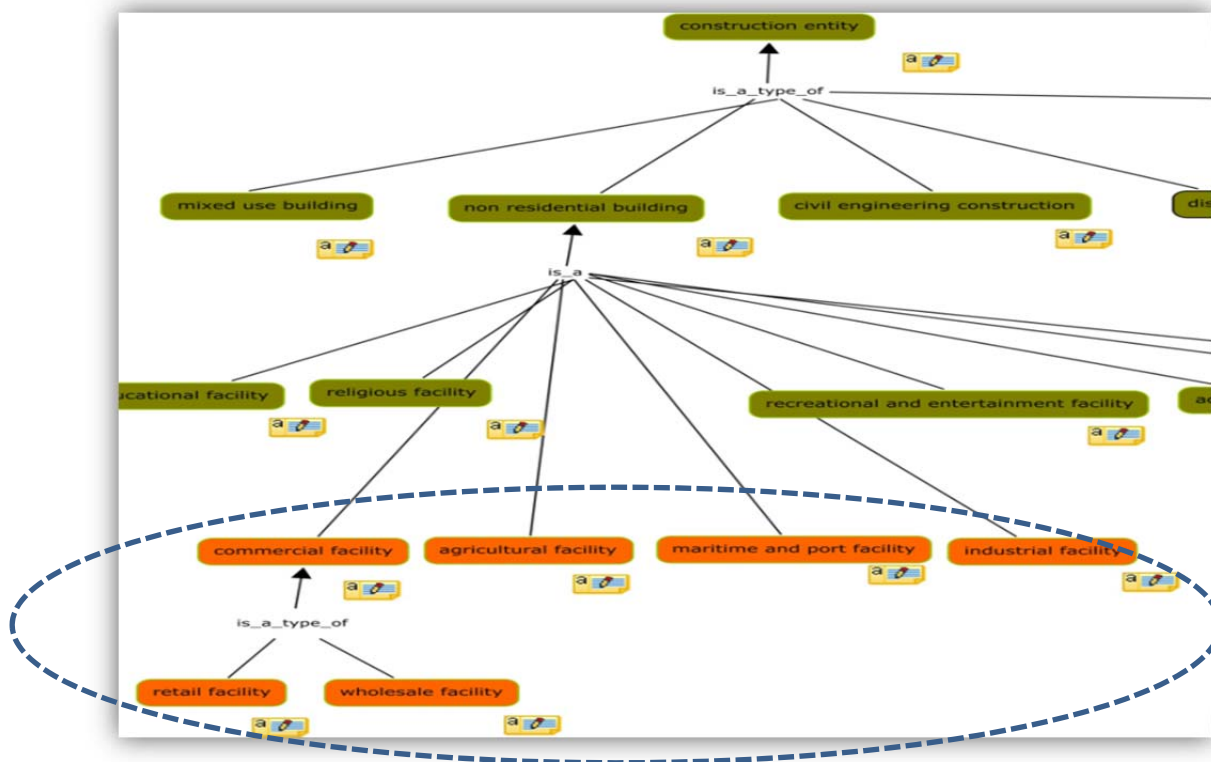


Fig. 26 - Proposta de adição de JF_{fr}

Esta adição foi justificada por JF_{fr} tendo em conta o tipo de público-alvo da Associação²⁰¹ a que pertence o especialista e as áreas e tipos de trabalho em que os seus associados desenvolvem a sua actividade e que não se encontravam, na sua perspectiva, totalmente representadas²⁰².

²⁰⁰ **Nota 1:** Os exemplos apresentados conterão os conceitos propostos pelos especialistas designados em língua inglesa, por uma questão de simplicidade e de homogeneidade. Como é natural, nos mapas conceptuais foram incluídos os equivalentes na língua de trabalho do especialista que sugeriu a sua inclusão.

Nota 2: A conceptualização de mapas construídos exclusivamente nas línguas de chegada foi sendo desenvolvida, de modo mais informal, ao longo deste processo e do de localização, tendo a sua análise sido útil sobretudo para exemplificar pontos de vista e analisar opções. Estas reflexões foram, posteriormente, reflectidas nos mapas conceptuais, através de propostas de reconceptualização que incluíam já a identificação dos equivalentes.

²⁰¹ CAPEB - Confédération de l'artisanat et des Petites Entreprises du Bâtiment - <http://www.capeb-paris.fr/>

²⁰² Exemplo de uma necessidade idêntica foi também apresentado por um especialista de outro país. No seu comentário a este mapa JPM_{pt} afirma: "Uma vez que se está a dividir os edifícios em muitas subcategorias, talvez se devesse fazer o mesmo para os outros trabalhos de engenharia (estradas, pontes, infra-estruturas, etc.). Não conheço outra classificação para os edifícios que seja mais aplicável ao caso nacional".

Exemplo 2:

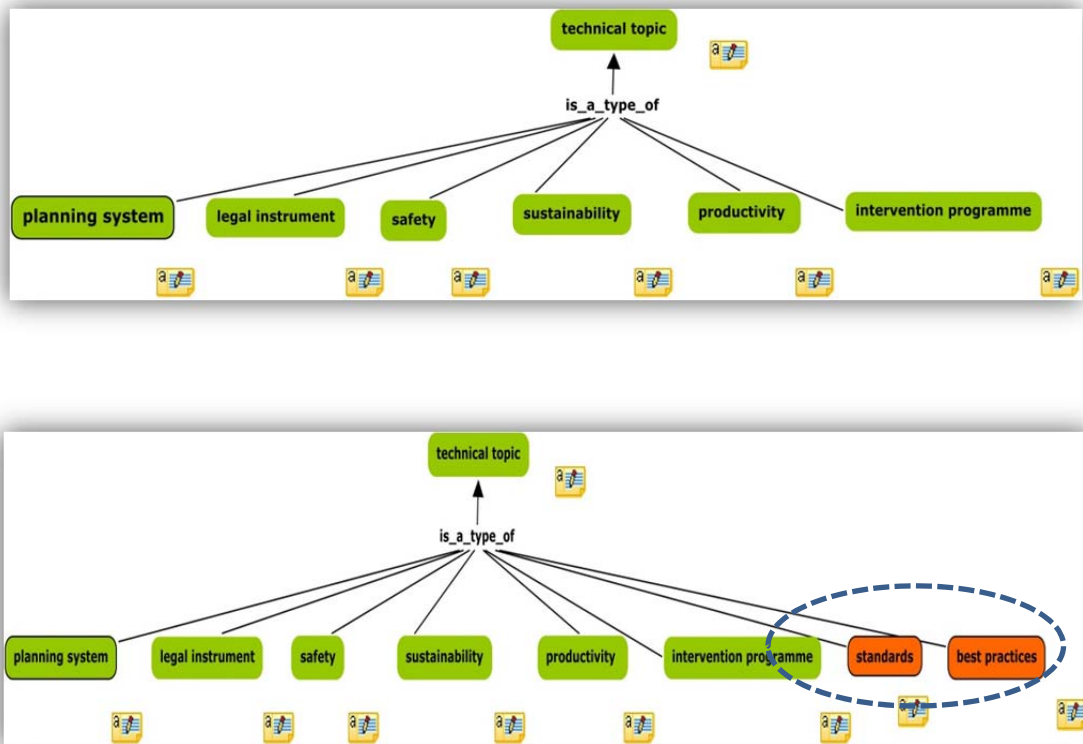


Fig. 27 - Proposta de adição de JPM_{pt}

No seu comentário²⁰³ ao mapa *technical_topic*, JPM_{pt} afirma: “este mapa contém temas muito díspares e que não definem de forma completa o tópico (penso que isso seria impossível neste caso). Sugiro que seja acrescentado um nó “Normas” (sem força de lei) e “Regras da boa prática” que, em Portugal são muitas vezes referidas na ausência de uma base normativa e/ou legal adequada”.

²⁰³ Os comentários que aqui transcrevemos foram obtidos a partir dos comentários inseridos no guião, no espaço concebido para esse efeito.

Exemplo 3:

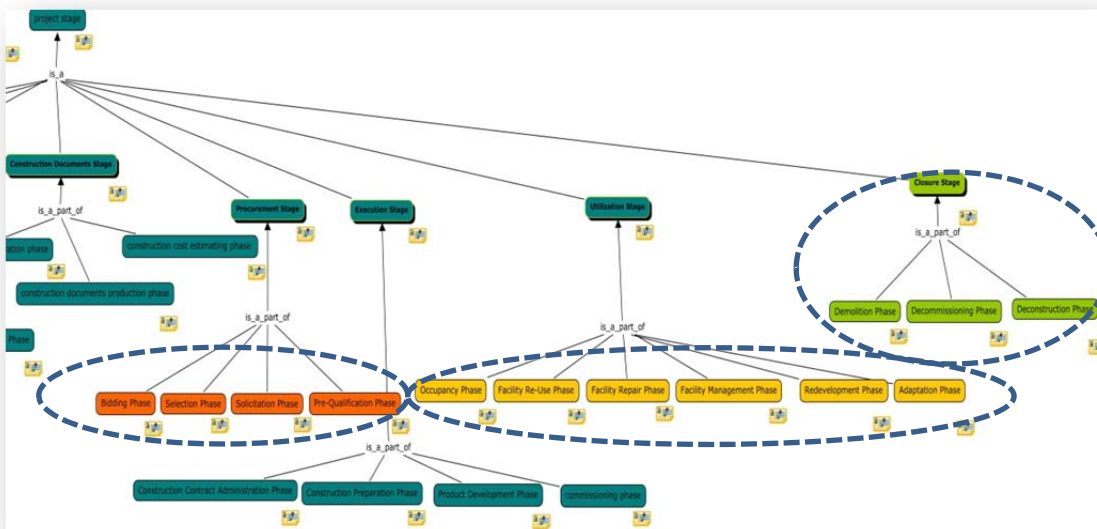
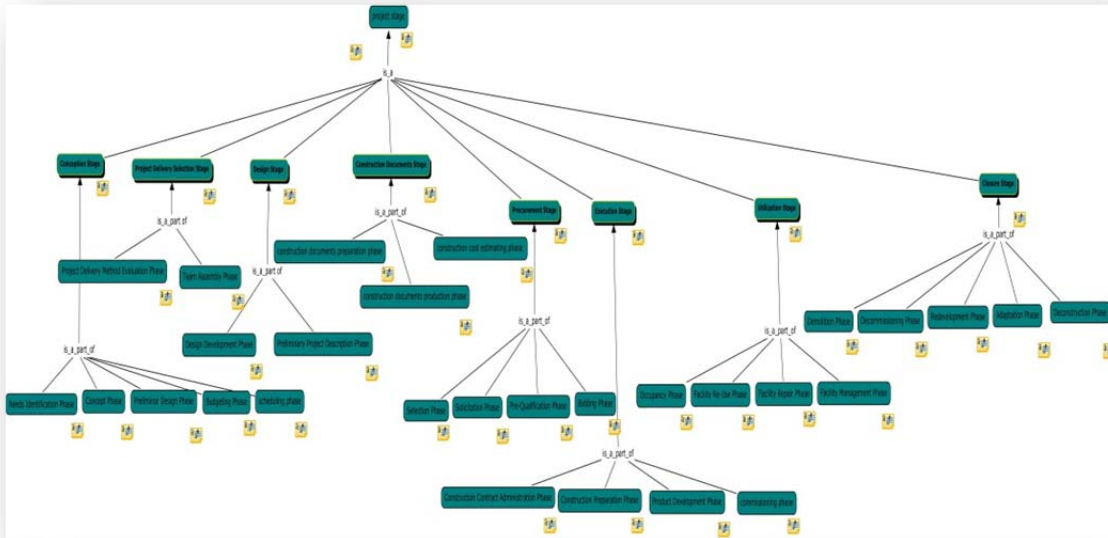


Fig. 28 - Proposta de modificação de AS_{pt}

No seu comentário AS_{pt} afirma que: (...) “*Selection deve estar depois do bidding; As fases iniciais dos empreendimentos estão muito detalhadas para o que é habitual; (...) Redevelopment e adaptation não deveriam estar na utilization phase*”.

Exemplo 4:

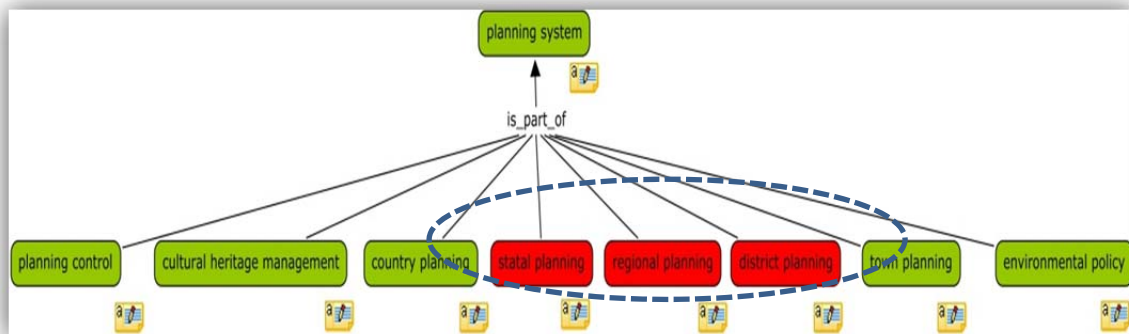


Fig. 29 - Proposta de adição de GH_{it}

Um outro exemplo de proposta de adição foi apresentado pelo especialista italiano GH_{it}, que sugeriu a adição de três conceitos no mapa *planning_system*, o que corresponde à sua visão do conjunto de entidades com capacidade para interferir no processo de planejamento em termos da realidade italiana.

- b. adição**, resultante da necessidade de **representar o conceito de partida em mais do que um conceito na língua de chegada**.

Este processo, para além de representar novos conceitos, pode implicar também a possível alteração e a redefinição da posição relativa dos já representados, bem como das relações existentes entre os conceitos, como se ilustra na proposta seguinte:

Exemplo 1:

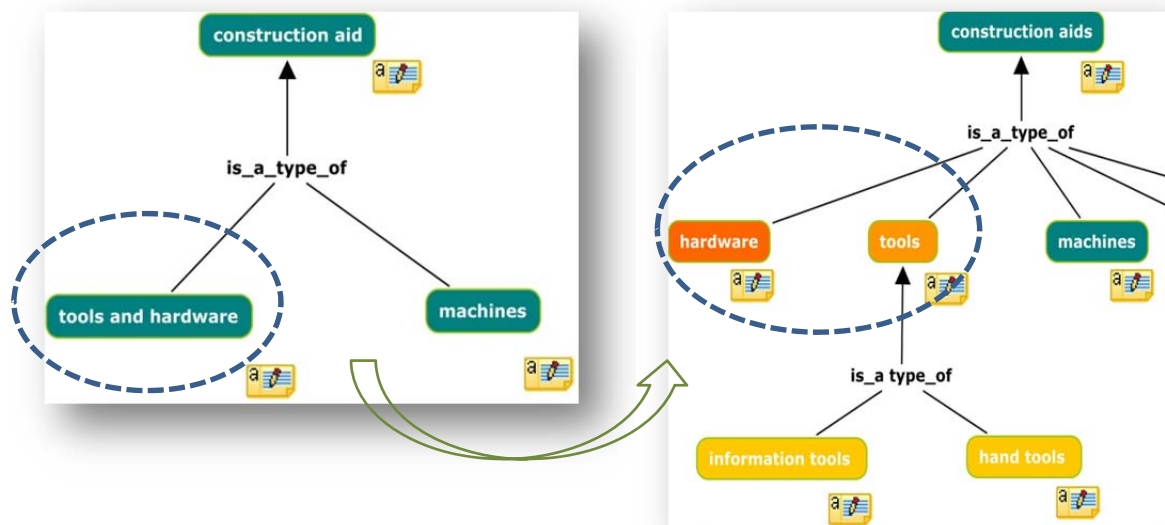


Fig. 30 - Proposta de adição/modificação de JF_{fr}

Para JF_{fr} esta modificação é essencial para poder estabelecer uma distinção entre ferramentas, sobretudo ao nível do tipo de ferramentas de apoio à construção

Como descrevemos antes, os mapas conceptuais constituintes da amostra apresentam diferentes níveis de granularidade e complexidade, como se pode observar pelo Apêndice H. No entanto, como pudemos verificar, o fenómeno da adição (e divisão) não se verificou exclusivamente em mapas com menor grau de granularidade ou menor número de conceitos, o que nos poderia levar a concluir, erroneamente, que as representações conceptuais de mais alto nível seriam mais permeáveis a este fenómeno, porque eventualmente incompletas no que respeita à representação do conhecimento do domínio do ponto de vista do especialista.

- 2. Modificação da representação conceptual.** Neste processo os conceitos mantêm-se constantes, mas pela análise da representação existente, os especialistas introduzem novas perspectivas, conducentes à redefinição da sua posição inicial e, eventualmente, das relações conceptuais que estabelecem com os restantes conceitos.

Exemplo 1:

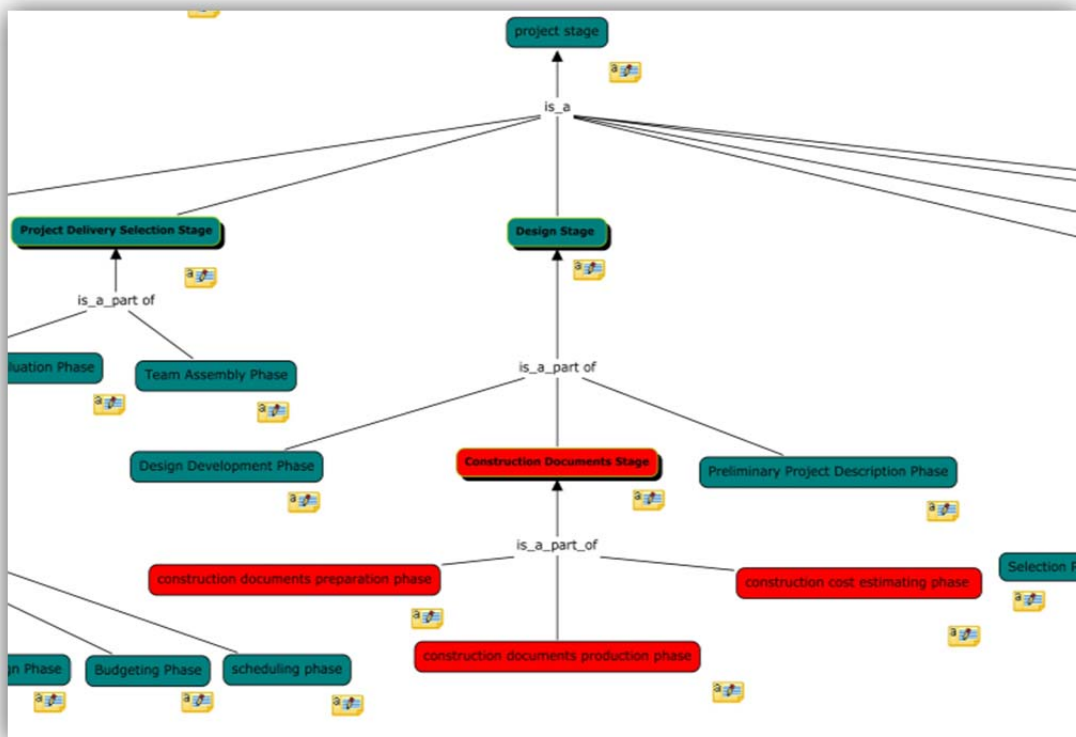
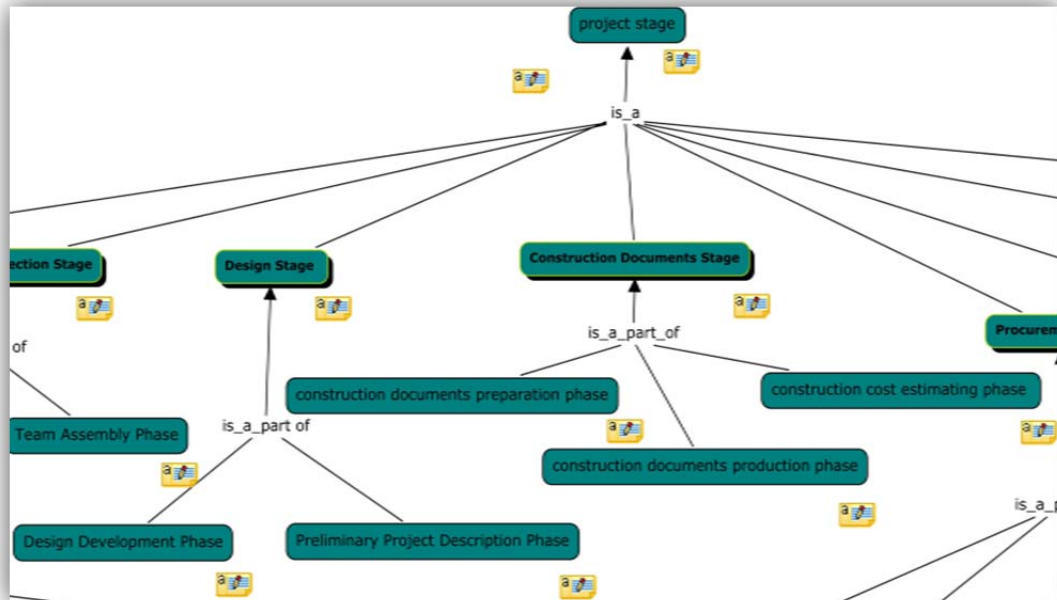


Fig. 31 - Proposta de modificação de JPM_{pt} e de AS_{pt}

Segundo o comentário de JPM_{pt}: “o processo construtivo tem menos fases e a fase de elaboração da documentação surge associada, normalmente, à de design”. Já para AS_{pt}: (...) “A preparação e elaboração dos documentos de concurso devem ser na mesma fase!”.²⁰⁴

A coincidência na opinião expressa pelos especialistas anteriores, quanto à necessidade de redução do número de fases do processo construtivo partiu de uma análise da sua realidade nacional, mas foi adoptada como uma perspectiva coerente com as restantes realidades pelo conjunto dos especialistas.

Finalmente, no que respeita aos fenómenos observados que não alteraram a representação conceptual em termos do número de conceitos e da sua estruturação, estes referem-se sobretudo à existência de termos que podem ser alvo de uma modificação semântica, *i.e.*, em que o termo de partida é localizado para um termo na língua de chegada, que tem um significado mais específico ou um significado mais genérico, mas mantém a sua posição na representação.

3. **Eliminação de conceitos.** Ao longo do desenvolvimento do processo de localização os especialistas propuseram, após um processo de revisão, a eliminação de alguns conceitos. Esta eliminação foi proposta para um número diminuto de conceitos por entenderem que:

- a.** os conceitos (1) não se adequam à representação, ou (2) podem/devem ser representados noutros mapas onde a sua presença seria mais lógica ou, finalmente, (3) não representam conhecimento específico do domínio.

²⁰⁴ Neste ponto podemos identificar uma concordância entre os dois especialistas de origem portuguesa.

Exemplo 1:

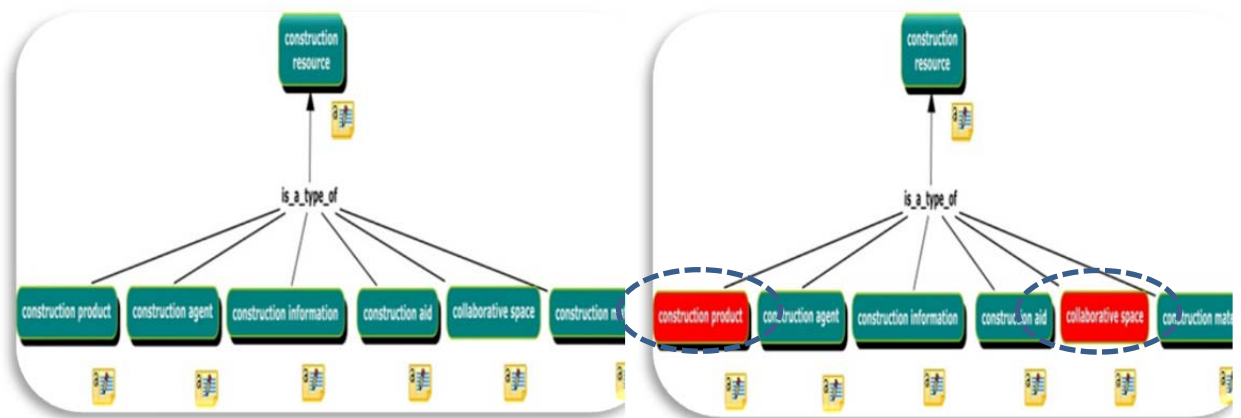


Fig. 32 - Proposta de eliminação de JPM_{pt}

JPM_{pt}- no seu comentário ao mapa *construction_resource*, afirma: “Eliminaria “Construction Product” (eventualmente incluiria esses elementos como Materiais). Eliminaria também os “espaços colaborativos”, que colocaria sob uma nova categoria (“Sistemas de Informação para a Construção”). Tenho aqui uma dúvida: faz sentido abrir uma nova categoria aqui para serviços (águas, electricidade, internet, etc.?)”.

Para LJU_{de} o conceito: “*collaborative space*” is not construction industry specific term and should not be included in the corresponding ontology – it was somehow “forcedly invented” in H-KNOW. Esta opinião é coincidente com a de AS_{pt}, que afirma: “Não sei se *collaborative space* será resource. Só se for no âmbito do H-KNOW”.

O conceito de *collaborative_space*, apesar de representar, no âmbito da plataforma H-Know, um recurso de auxílio à partilha do conhecimento do domínio e, por consequência ao desenvolvimento do processo construtivo, foi, como podemos perceber pelos comentários anteriores, o que colocou maiores dúvidas aos diversos especialistas, dada a sua grande especificidade e dependência da visão e do sistema desenvolvido no âmbito de um projecto concreto.

- b. Por entenderem que os conceitos não existem na sua realidade de trabalho / na sua língua de trabalho.

Exemplo 1:

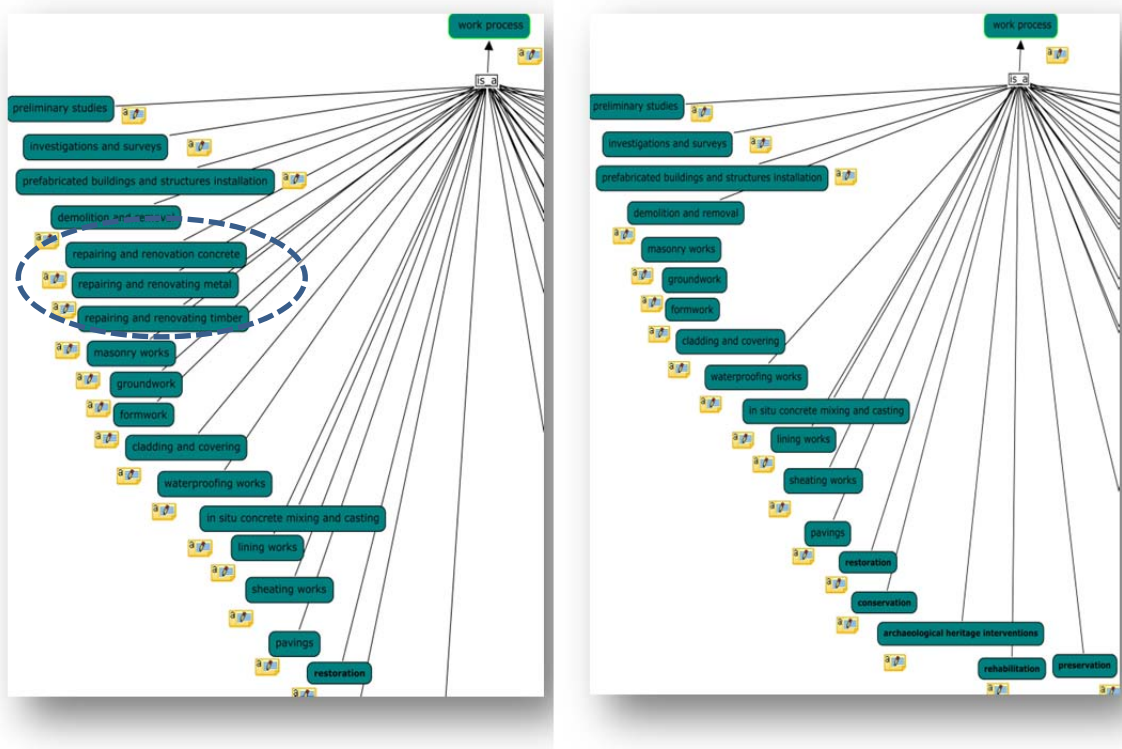


Fig. 33 - Proposta de eliminação de AS_{pt}

AS_{pt} explicita que propõem a eliminação dos conceitos “*repairing and renovating concrete, repairing and renovating metal e repairing and renovating timber*”, uma vez que, na sua perspectiva: “*em PT, renovação do betão, metal e madeira não existe*”.

As propostas anteriormente descritas constituem exemplos que contribuem para demonstrar que os problemas específicos de localização de uma ontologia não se situam só ao nível linguístico (morfológico, gramatical ou sintático), advindo também da necessidade de considerar, em paralelo, os aspectos conceptuais próprios ao domínio e à representação do conhecimento em cada língua de chegada, o que pode ser facilitado, no nosso ponto de vista, pelo desenvolvimento iterativo e simultâneo dos processos de conceptualização e localização e pela presença do especialista.

Como explicita (JPM_{pt}) na sua resposta à pergunta catorze da entrevista:

O especialista se estiver a conceptualizar pensa mais no seu contexto – concentra-se nele, recorre à sua formação e à sua experiência. Quando está em presença de mais do que uma língua é obrigado a concentrar-se também nos pontos comuns da representação – sobretudo os de mais alto nível. A presença das diferentes línguas suscita reflexões sobre os diferentes contextos (países) e coloca o especialista perante outros desafios.

Esta afirmação, reforçada pelos pontos de vista de outros especialistas, que explicitaremos no ponto seguinte, deixa-nos perceber que estes, ao seguirem a abordagem proposta, de base conceptual, não fizeram uma comparação simples dos termos entre si, mas tiveram em conta o sistema conceptual que estava na base do uso dos termos. Por outro lado, quando nos deparamos, como acontece no exemplo anterior, com conceitos possuidores de uma carga cultural e profissional específica, torna-se importante ter presente e recordar/confrontar o especialista com os propósitos de construção da ontologia, de modo a que este coloque em perspectiva as suas decisões.

c) a sugestão de **eliminação de conceitos** passa, em alguns casos, pela proposta de substituição da representação conceptual por uma mais próxima da realidade do especialista, tal como se percebe pelo comentário abaixo:

JPM_{pt}, quando reanalisa o mapa *work process*, afirma: *“Fiz várias sugestões, mas creio que para a realidade nacional seria preferível substituir integralmente este mapa por um outro baseado na classificação usada pelo ProNIC ou, eventualmente, nas regras de medição do LNEC. Esta classificação, mais do que a referida no mapa anterior, depende MUITO das normas nacionais e da cultura local”.*

Este exemplo permite-nos perceber a necessidade de adaptação dos recursos (mencionada no ponto 3.3.1). De facto, os diferentes recursos semânticos que o especialista considera ou a que recorre estão, normalmente, em diferentes estádios de desenvolvimento, sendo que constatámos, como afirmam Hubrich *et al.* (2008: 266), que a estrutura e a especificidade dos recursos semânticos locais se torna mais evidente neste processo de comparação com recursos internacionais.

Finalmente, importa destacar que a análise conceptual do especialista leva-o a confrontar-se com domínios onde se sente menos confortável a analisar a representação do conhecimento, uma vez que esta, graças às suas características de interdisciplinaridade ou à sua especificidade, pode, na perspectiva do especialista, requerer a presença ou a consulta a especialistas de outras áreas, como se pode perceber pelo comentário seguinte, em que JPM_{pt}, a propósito do mapa *planning_system*, refere:

Este mapa procura caracterizar actividades na área do planeamento, pelo que não estou inteiramente familiarizado com estes tópicos. Ainda assim, penso que “ordenamento do território” (disciplina) não admite como variante “planos de ordenamento do território” (instrumentos). Admito que estes conceitos sejam variáveis de país para país, em função dos princípios que presidam às orientações de planeamento (ambiente, tipo de uso, volumetria, cêrcea, etc.). (...) Sugiro que este tópico seja devidamente desenvolvido, com o apoio de um especialista na área do planeamento.

Já no que respeita à **análise e utilização das relações conceptuais**, não foram propostas, por parte dos especialistas, alterações substanciais às relações já existentes nas representações conceptuais, tendo, no entanto, sido sugerida:

1. a eliminação de relações conceptuais:

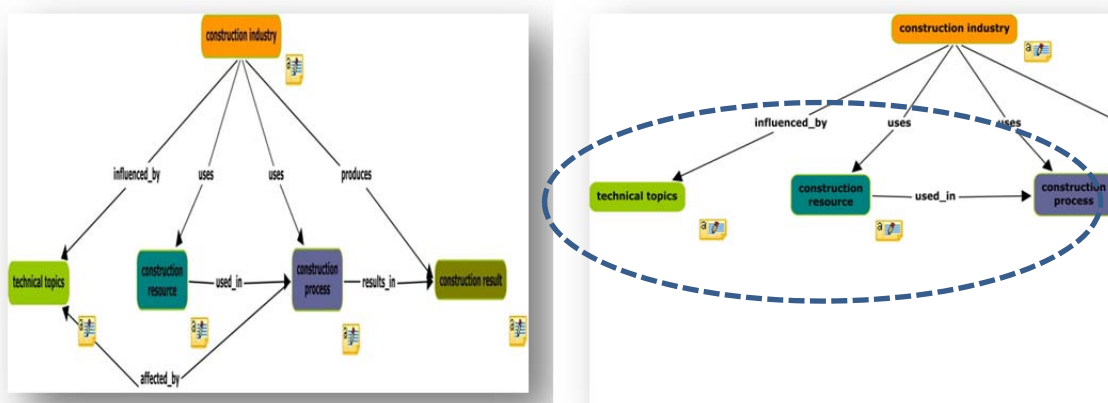


Fig. 34 - Proposta de eliminação de JPM_{pt}

JPM_{pt}, no mapa *Global_frame_of_the_high_level_ontology*, sugere: (...) “uma variante para “Resultados da Construção”. Sugiro eliminar a relação entre “Tópicos Técnicos” e “Processos de Construção” (...)”.

2. a adição de relações conceptuais:

- a. adição de relações do tipo associativas (relação de espaço – *takes_place*), que surge associada sobretudo à representação (descrição) de processos construtivos típicos do domínio.

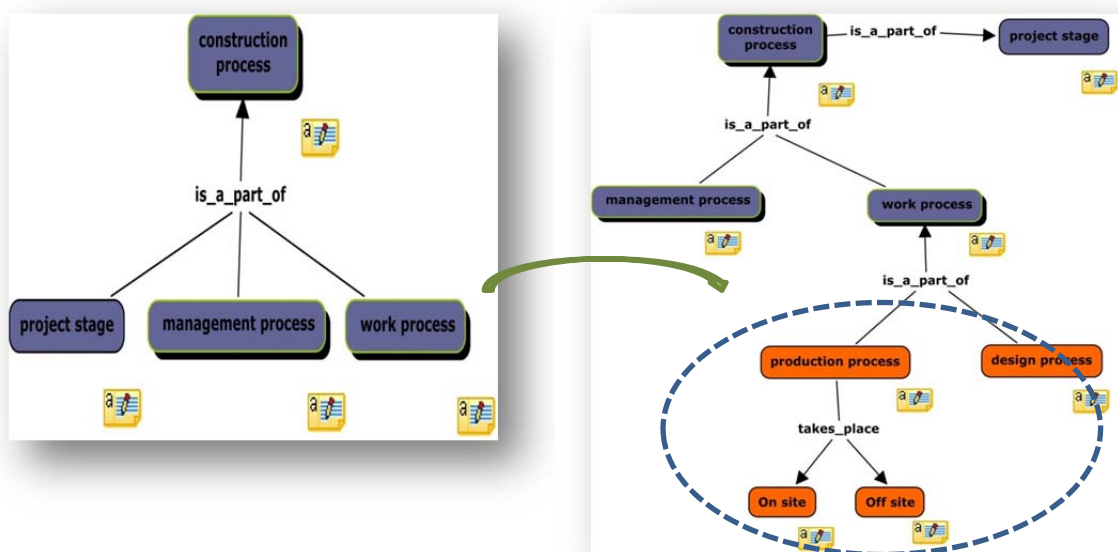


Fig. 35 - Proposta de adição de JPM_{pt}

Pudemos observar que a relação mais usada nas propostas de reconceptualização foi a relação hierárquica do tipo partitivo²⁰⁵ – *is_a_type_of*. O uso da relação partitiva, transversal a todos os domínios do conhecimento, é recorrente nos processos de conceptualização antes identificados e que são mais usuais: (1) os especialistas reconheceram lacunas e acrescentaram, a partir do seu ponto de vista, conceitos em falta; ou (2) acrescentaram aos conceitos genéricos conceitos mais específicos que, mais uma vez, correspondem a lacunas/necessidades identificadas em função da sua realidade profissional e cultural.

²⁰⁵ Usamos a classificação proposta na norma ISO 704-EN:2009.

A escolha e a inserção de relações conceptuais causaram, como observámos, algumas dúvidas nos especialistas, dada a sua formação escassa no uso deste tipo de representação conceptual e o seu algum desconhecimento quanto aos diferentes tipos de relações existentes. Para a resolução destas dificuldades torna-se importante a presença de um terminólogo, ou de um gestor de conhecimento, para esclarecer eventuais dúvidas e exemplificar quais as possibilidades de uso das diferentes relações conceptuais.

A falta de esclarecimento quanto ao uso destes elementos pode tornar a participação do especialista menos efectiva e dificultar a obtenção dos seus contributos. Por outro lado, o número limitado de alterações, bem como o uso de relações iguais às já existentes nos mapas, permitiu uma integração mais fácil entre as ontologias das diferentes línguas, facilitando a sua implementação formal e beneficiando a sua interoperabilidade.

4.2.5.2. Análise prévia dos resultados

Os resultados obtidos a partir da experimentação do método permitem-nos concluir que nos processos de análise da conceptualização inicial e da sua localização, a escolha e definição dos equivalentes foram efectuadas pelos especialistas em função do valor que os conceitos, e os termos que os designam, assumem na representação conceptual, funcionando a língua como ponto de acesso ao equivalente que melhor se adequava ou que melhor designava o conceito na língua de trabalho do especialista.

A integração entre o nível conceptual e o nível linguístico nem sempre foi fácil, como se percebe pelos exemplos apontados, mas observámos que a análise e resolução das dificuldades pelos especialistas deram origem a contributos relevantes para a análise conceptual que, de outro modo, poderiam não ter sido considerados.

No desenvolvimento do processo de análise da conceptualização inicial ficou claro que os especialistas se socorreram, em primeira instância, da sua formação, da sua experiência e dos instrumentos de organização do conhecimento do domínio que lhes são mais familiares, os quais estiveram na base também de algumas das propostas de reconceptualização.

O recurso ao conhecimento mais presente, normalmente próximo do da realidade nacional do especialista, justifica, na nossa perspectiva, a existência de um *corpus*, quer na língua de trabalho da rede colaborativa quer nas restantes línguas, que sirva de auxílio ao processo de (re)conceptualização e de localização, necessidade reconhecida pelos especialistas ao introduzirem novas fontes ao longo do processo. Esta possibilidade deve, no entanto, ter em consideração e ser adequada aos requisitos, limitações e às necessidades inerentes ao funcionamento da rede.

Observámos, também, que a explicitação das dúvidas durante o processo de localização pode conduzir à tentativa de conjugação dos pontos de vista, quer entre especialistas quer entre o conhecimento específico de cada um e a organização do conhecimento com que é confrontado, por exemplo, ao considerar e reutilizar normas ou classificações. Esta tendência para a conjugação acontece porque o especialista se revê como elemento de uma rede colaborativa, onde contribui para a construção de uma representação conceptual do seu domínio de conhecimento específico, representação que vai para além do que seria a sua representação particular desse conhecimento, valorizando, assim, o conjunto das opiniões e dos saberes.

De facto, no processo construção de uma representação multilingue do conhecimento do domínio, torna-se inevitável que os participantes se separem dos seus pontos de vista específicos e procurem definir representações comuns. A localização e, posteriormente, a integração das representações das diferentes línguas podem, assim, tornar-se processos úteis na detecção daqueles elementos, sobretudo na construção de uma ontologia de domínio de alto nível.

Concluimos, igualmente, que o espaço de conceptualização e localização disponibilizado aos especialistas para desenvolver estes processos tornou o processo ágil e motivou a participação daqueles, tendo a utilização de mapas conceptuais, como ferramenta de apoio à representação do conhecimento, facilitado a implementação do método. Concluimos, ainda, que esta forma de representação é capaz de sustentar e promover a participação eficiente dos especialistas no processo de eliciação e organização do conhecimento especializado e de promover o trabalho interdisciplinar, nomeadamente com o terminólogo.

Para este efeito contribuiu também a definição prévia de procedimentos e ferramentas a utilizar para apoiar o desenvolvimento de ambos os processos, descritas no âmbito do guião elaborado para os intervenientes na experiência, bem como o acompanhamento do desenrolar desta pelo terminólogo.

A análise dos dados obtidos permitiu-nos, ainda, observar que a estruturação do sistema conceptual pode ser, até um determinado nível e de acordo com propósitos bem estabelecidos, dinâmica e flexível, de forma a poder reflectir os resultados das propostas de reconceptualização decorrentes da participação dos especialistas neste processo.

Um desafio especial colocou-se, no entanto, na necessidade de distinguir claramente entre a construção da conceptualização de alto nível e a vontade de elaboração ou de uso de representações conceptuais mais específicas (locais), que não se enquadravam no âmbito e nos propósitos da ontologia de alto nível H-Know. Esta preocupação foi sendo reiterada e explicada de modo a evitar a criação de um grande número de representações adicionais e a proliferação de conceitos e com o intuito de obter um contributo dos diferentes especialistas que se enquadrasse no âmbito e nos propósitos definidos para a construção da ontologia.

Em resultado deste processo e dos seus contributos, as propostas de reconceptualização foram posteriormente analisadas e revistas, tendo apenas uma parte sido considerada e incluída na versão final da ontologia. Estas decisões foram tomadas em conjunto pelos especialistas portugueses, responsáveis pelo acompanhamento do processo de construção da ontologia no âmbito das tarefas do projecto, e pelo terminólogo, tendo em conta os objectivos definidos para a construção da ontologia de alto nível H-Know.

As decisões tiveram em conta a necessidade de dotar os utilizadores do H-Know e da sua plataforma colaborativa, de uma perspectiva e de uma representação comuns do conhecimento do domínio da reabilitação, que suportasse os serviços relacionados com a monitorização do conhecimento e com a extracção de informação e do conhecimento específico ao domínio, potenciando, assim, as actividades da rede colaborativa e os dos seus utilizadores.

Neste processo o terminólogo conjugou as diferentes propostas e a diversa informação disponível, actualizou as representações conceptuais e desenvolveu um papel de

mediação e de apoio à decisão, tendo, posteriormente, acompanhado o desenvolvimento da implementação formal da ontologia.

A experiência levado a cabo permitiu-nos observar que o método proposto pode ser utilizado de modo eficaz no desenvolvimento do processo de localização e na especificação multilingue da ontologia. Por outro lado oferece um fluxo e um espaço de trabalho que sustenta também a fase de reanálise da conceptualização do domínio e a participação dos especialistas no seu desenvolvimento.

Permite, finalmente, que se considerem, para as eventuais decisões de reconceptualização, os diferentes pontos de vista e as diferentes realidades nacionais, apontados pelos especialistas ao longo do processo, conduzindo assim, a uma construção efectivamente partilhada de uma ontologia representada em mais do que uma língua natural, cujo objectivo é o de facilitar o entendimento mútuo, ultrapassando as barreiras impostas pelas línguas.

4.2.5.3. Inquérito aos especialistas

Como parte das actividades desenvolvidas para a análise e validação dos dados obtidos durante a realização da experiência, concebemos um inquérito, sob a forma de entrevista semiestruturada. A entrevista, bem como a sua tradução podem ser consultadas nos Apêndices D e E, respectivamente²⁰⁶. A entrevista foi elaborada com o objectivo de:

1. analisar a opinião dos especialistas quanto ao contributo e validade do método proposto para a:
 - a) estruturação e representação do conhecimento do domínio,
 - b) localização das representações conceptuais,
 - c) partilha de conhecimento e obtenção de consenso,
 - d) qualidade final da ontologia.

²⁰⁶ A totalidade das transcrições obtidas nas entrevistas está disponível para consulta no Apêndice F.

2. inquirir os especialistas da rede colaborativa quanto à percepção que têm sobre o papel que desempenharam na (re)conceptualização do conhecimento do domínio e no processo de localização,
3. identificar as necessidades e dificuldades dos especialistas, quer no desenvolvimento do processo de localização quer no uso do método proposto,
4. identificar as suas propostas para a melhoria do método.

A realização da entrevista aos especialistas envolvidos na experiência funcionou como um complemento à restante observação e recolha de dados e como um meio para explorar e validar os resultados alcançados e analisar as opiniões dos participantes, tendo sido inquiridos cinco especialistas de quatro dos países participantes²⁰⁷.

A entrevista semiestruturada foi dividida em três partes, com a intenção de obter:

1. uma breve caracterização de cada especialista, também no que respeitava ao conhecimento e uso de recursos de tradução e de localização, de modo a perceber a maior ou menor facilidade em levar a cabo o processo de localização,
2. *feedback* sobre o desenvolvimento dos processos de (re)conceptualização e localização, e
3. *feedback* sobre a aplicação do método.

Uma parte das respostas obtidas e dos pontos de vista expressos foi já considerada e analisada em pontos anteriores deste capítulo. Pretendemos, neste ponto, analisar as questões remanescentes, ligadas sobretudo à aplicação do método e à percepção que os especialistas retiveram quanto ao papel que desempenharam.

Começaremos assim a nossa análise pela questão número seis. Esta questão destinava-se a perceber se, no contexto de construção de uma ontologia, os especialistas consideravam que podiam desempenhar um papel activo no encontro e na escolha dos equivalentes para a sua língua. A resposta obtida foi positiva em todos os inquiridos, tendo estes justificado o seu ponto de vista com diversas razões, ligadas normalmente à formação e experiência acumuladas.

²⁰⁷ Não obtivemos resposta à entrevista por parte dos especialistas italianos.

JPM_{pt} afirma que terá que mesmo que ser o especialista a desempenhar esse papel, dado que uma *“ontologia não é um objecto puramente linguístico”*, requerendo o seu significado *“uma compreensão e uma explicação extensa”*. AA_{es} reforça essa perspectiva ao afirmar que os especialistas possuem *“a holistic perspective of the domain”*, e acrescenta que estes estão conscientes do seu papel dado que conhecem os diversos *“points of view of the organizations involved in the construction of the ontology and acknowledge the need to reach a common view of the domain”*.

Estas afirmações vão de encontro ao nosso ponto de vista, expresso no ponto 3.8., em que descrevemos os especialistas como actores activos e conscientes de um processo onde, em paralelo, são produtores e utilizadores de um artefacto semântico multilingue, artefacto que representa o conhecimento do domínio e que é desenvolvido de acordo com um conjunto de critérios e propósitos bem estabelecidos, no contexto de um projecto e de uma rede e das limitações socio-funcionais e temporais inerentes à sua construção, especificação formal e implementação.

Os especialistas possuem, de facto, e como pudemos observar, a capacidade de providenciar dados, intuição, análise da qualidade e análise crítica ao longo de todo o processo, sendo que o modo como percebem a necessidade de estruturação do conhecimento do domínio está, normalmente, mais ligado aos aspectos contextuais e funcionais do domínio e da própria rede e não tanto à necessidade de localização *per se*. Foi, no entanto, o desenvolvimento desta, que suscitou, como reconheceram, uma parte das reflexões e propostas de reconceptualização.

As questões nove a onze destinavam-se a obter *feedback* sobre os elementos disponibilizados pelo método para a realização dos processos de conceptualização e de localização.

Assim, quando questionados²⁰⁸ sobre se os passos e elementos constantes dos mapas conceptuais – equivalentes, variantes, definições (e identificação das fontes), contexto de uso – eram relevantes e auxiliavam à representação do conhecimento e à sua localização nesta fase de especificação semiformal, responderam na sua totalidade que sim, tendo AS_{pt} acrescentado que: *“ajudavam a definir o conceito. As palavras não possuem um*

²⁰⁸ Questão número dez.

só sentido. Assim delimita-se o uso da palavra e enquadra-se melhor o conceito que se está a analisar”.

AA_{es}, por sua vez, referiu que *“it helps you to start the process and helps to achieve a common understanding, as you can investigate the concept through the sources and also find the translations there”*. LJU_{de}, por seu lado, considera necessários outros elementos para outras fases²⁰⁹ mais formais da representação do conhecimento: *“they offer enough information for the basic level (semiformal), not really for the advanced one - level of higher expertise (where more relationships are needed)”*.

Questionados²¹⁰ sobre se, na sua perspectiva, os mapas conceptuais seriam uma ferramenta viável para a representação do conhecimento de um domínio específico em mais do que uma língua, a resposta que obtivemos foi unânime, tendo AS_{pt} destacado que funcionam sobretudo nesse aspecto, dado que: *“permitem ao especialista interpretar a realidade representada e a tradução permite que se relacione/se ponha em perspectiva a interação entre os processos e entre as diferentes fases (relativas ao domínio)”*.

JF_{fr} destacou, em resposta à questão doze, a importância do seu uso em comparação com os recursos que recorrem a outras formas de representação, como a hierárquica, e afirmou que:

“It is a good tool and clear to see the complexity (...). It’s easier to use conceptual maps and to understand them, than to use a classification like Omniclass - due to the use of elements like relations. Omniclass is like a catalog – you cannot have a global view of things - it is not possible. Maps have an advantage – the representation / the spatial division, which helps us to understand each concept”.

De modo avaliar o ambiente de conceptualização e localização proposto e a perceber se introduziriam alterações a este ambiente, questionámos²¹¹ os especialistas sobre que outros elementos considerariam relevantes para tornar mais eficiente o processo de conceptualização e localização com base em mapas conceptuais.

²⁰⁹ JPM_{pt} sugere a existência de diferentes fases neste processo. Para o especialista deveriam existir duas fases: *“Uma 1ª fase de conceptualização livre, do tipo folha de papel em branco, e uma 2ª fase mais restrita, guiada e com estes ou outros elementos. O processo deveria ser iterativo”*.

²¹⁰ Questão número nove.

²¹¹ Questão número onze.

Na maior parte dos casos não houve sugestões ou foi reconhecido que os elementos eram suficientes, uma vez que, como afirmou JF_{fr}: *“if the maps are too complex / have too many elements it won’t be easy to work. For localization and translation that would be no problem – for the conceptualization – it would be very difficult”*, acrescentando que não havia necessidade de *“complexify the complexity – systematically analyzing the maps is complex enough”*. Apenas AS_{pt} sugeriu a inclusão de figuras e gráficos, enquanto AA_{es} sugeriu a inclusão de *“examples of sentences involving the concepts”*, mas disponibilizados previamente ao especialista pelo terminólogo/tradutor.

Estas respostas e reflexões, acrescidas da observação dos resultados do trabalho feito com base neste ambiente, deixam-nos perceber que o espaço de conceptualização e localização, que propomos para apoiar o processo especificação multilingue de uma ontologia, no seio de uma rede colaborativa multilingue, é funcional e de uso acessível, uma vez que o espaço de trabalho permite, a utilizadores sem grande experiência em conceptualização e localização e que não estão preparados para lidar com as restrições da semântica formal, que se concentrem nas tarefas de análise e de localização da representação conceptual num ambiente menos restritivo.

As restantes questões, da catorze à dezassete, destinavam-se sobretudo a obter *feedback* quanto ao uso do método proposto para os processos de conceptualização e de localização.

Assim, questionámos²¹² os especialistas sobre se consideravam que a presença de mais do que uma língua poderia contribuir para atingir um maior nível de consistência na construção da representação conceptual do conhecimento relativo ao domínio e, em consequência, da ontologia.

Todos responderam positivamente, tendo evidenciado o papel complementar que a presença de mais do que uma língua pode proporcionar na identificação de lacunas e conceitos em falta, ou o modo como essa presença *“sometimes help to get precise information on the specific term – for somebody who can understand different languages”* (LJU_{de}). Por outro lado, esta presença é vista por AA_{es} como um elemento que *“leverages the common understanding of the concepts”*. Segundo este, a existência de:

²¹² Questão número catorze.

a consistent translation is a good thing for all. You assume there is consistency while talking about concepts. You achieve a common view of the different terms, which is very important to motivate participation and the use of the tool. It also helps in the communication and helps people to better know each other and the different points of view.

JF_{fr} analisa esta situação de um outro ponto de vista e destaca a necessidade de se iniciar o processo por uma única língua, comum, *“to develop the enrichment process. Too many languages may be complicated – more than complexity you may find yourself complicating the process”*. Vai um pouco além e afirma que *“We should never depart from the particular languages – we have to start with one model in one common language and adapt to the specificity of the country”*.

Na sequência da questão anterior, perguntámos aos especialistas se, na sua opinião, o processo de localização, desenvolvido na fase de especificação semiformal do conhecimento do domínio, poderia contribuir para uma maior partilha de conhecimento entre os especialistas do domínio, tendo a resposta sido positiva na totalidade dos casos.

Para AA_{es} este processo devia existir *“from the beginning, using a continuous improvement approach cycle to validate it”*. Para JPM_{pt} este processo conduz a uma intervenção *“mais substancial do especialista. No entanto o processo torna-se, por vezes, mais difícil e mais complexo”*, pelo que propôs o recurso a técnicas de negociação formais para apoiar esta fase. Para JF_{fr} este tipo de processo permite perceber que *“we haven’t got the same organization and we can have a better knowledge of the organization of each other, perceive the differences and have a better knowledge of what each one does and how”*.

Questionados, em seguida, sobre se consideravam que o processo de localização, desenvolvido na fase de especificação semiformal do conhecimento do domínio, poderia contribuir para a melhoria dessas representações e, em consequência, para um aumento da qualidade da ontologia, todos responderam afirmativamente, tendo AS_{pt} considerado que o processo conduzia a maior *granularidade, consistência e discussão* à volta das representações, visão partilhada por JF_{fr}, para quem o especialista se concentra *“in the quality of the map”*. Afirma mesmo que, por vezes, passou *“one hour to translate one concept”*, dado que a ontologia *“must be exact. It is not a middle thing – this process must*

take to the improvement of the ontology – it considers different points of view – which together can improve the ontology”.

Na última questão, perguntámos aos especialistas se foi a representação conceptual que influenciou as suas decisões de localização ou se foi a localização que potenciou a sua análise e visão da representação conceptual.

Para AS_{pt} este é um processo iterativo, não existindo *“uma resposta única. A representação inicial cria uma influência na localização. Depois da localização é preciso fazer uma análise e interpretação para ver se representamos o que lá estava”*. JF_{fr} partilha desta opinião, mas vai um pouco mais longe e afirma que existe uma interacção, uma vez que:

a conceptual representation influences the thought. Then, during the localization process there is a sort of interaction between the term and the map.²¹³ Of course you get more interference and more discussion with the presence of the maps. It is better for the brain than just a simple list – you have a vision of the whole, of the concepts and their relations.

Para JPM_{pt} a representação conceptual surge primeiro. *“No entanto o especialista parte do seu ponto de vista – um ponto de vista nacional/regional/local e a representação não se pode limitar à representação do que se conhece, pelo que o especialista precisa/deve ser confrontado pelo mediador com a segunda parte do processo para repensar e generalizar as suas ideias”*.

Estas respostas permitem-nos concluir que a abordagem proposta, de base conceptual, afecta o processo de localização levado a cabo pelos especialistas, sendo que a consecutividade proposta no método entre conceptualização e localização, e a iteração que gera, permitem a obtenção de contributos interessantes para a procura de um entendimento comum dos termos, para o processo de escolha dos equivalentes adequados, promovendo, por outro lado, uma análise mais aprofundada do conhecimento do domínio, o que pode resultar numa construção mais consistente da ontologia.

²¹³ Nosso sublinhado.

4.2.6. Fase 6 - Reestruturação do método

Para tornar a aplicação deste método viável, criámos um espaço de conceptualização e localização multilingue baseado em mapas conceptuais. Como se pode observar pelos exemplos apresentados, os especialistas fizeram uso dos diferentes elementos disponíveis no ambiente de conceptualização e localização para elicitar e transmitir o seu conhecimento e as suas propostas quanto à organização e representação deste.

O uso dos diversos elementos disponibilizados e a riqueza da informação que permitiram recolher, levou-nos, assim, a considerar que a sua existência contribui para enriquecer quer o processo de partilha de conhecimento sobre o domínio entre os membros de uma rede colaborativa quer o processo de especificação multilingue de uma ontologia de domínio no contexto de uma rede colaborativa.

No entanto, constatámos também, tal como aconteceu em trabalhos anteriores (Kharatmal & Nagarjuna, 2010; Pereira, 2010), que esta forma de representação de conhecimento baseada em mapas conceptuais tem um grande grau de complexidade, que resulta sobretudo de três aspectos: o significado dado aos termos por parte de cada comunidade, as expressões usadas para designar as relações entre conceitos, e o quadro epistemológico da comunidade (Kharatmal & Nagarjuna, 2010).

Esta complexidade tem tendência a aumentar, como pudemos observar, quando recorremos a mapas conceptuais para desenvolver a representação multilingue do conhecimento de um domínio de especialidade, o que pode resultar, por um lado, numa limitação na compreensão do método proposto e dos seus objectivos e, por outro, numa limitação ao bom entendimento das actividades a desenvolver. Torna-se, por isso, essencial definir e explicitar os passos a seguir e acompanhar o processo e os seus actores ao longo de todo o seu desenvolvimento, uma vez que, tal como afirma JF_{fr} em resposta à pergunta número cinco: *“The specialist is not an expert in ontology construction or in translation - he assumes another role - facilitates the translation and acts as an intermediary”*.

A observação da aplicação e experimentação do método proposto e as opiniões obtidas quanto à sua validade levaram-nos, assim, a equacionar a possibilidade de introduzir modificações no método que pudessem resultar em melhorias no fluxo de trabalho proposto

e na sua implementação futura. Para tal, consideramos o uso que os especialistas fizeram dos diferentes elementos disponibilizados, nomeadamente no ambiente de trabalho em CmapTools, tendo ponderado ainda as respostas dadas quando questionados sobre se os passos e elementos constantes dos mapas conceptuais eram relevantes e auxiliavam à representação do conhecimento e à sua localização ou se sugeriam outros.

Verificamos, nesta análise, que o elemento disponibilizado para os especialistas indicarem contextos de uso do termo - *usageContext* – apenas foi usado em três conceitos, identificados abaixo.

AS_{pt} Mapa *construction_aid*

construction_aid

hasTranslation: Meios de apoio na reabilitação

hasDefinition: Equipamentos, máquinas e ferramentas de apoio à reabilitação de edifícios

usageContext: Utilização temporária como suporte das tarefas

machines

hasTranslation: Máquinas

hasDefinition: Equipamento mecânico necessário à execução das tarefas

usageContext: Nas tarefas aonde sejam necessárias máquinas

JPM_{pt} Mapa *work_process*

work_process

hasTranslation: processo de trabalho

usageContext: fase de execução

Como se pode observar, e apesar de esclarecido qual o tipo de *input* que se pretendia recolher com o uso de *usageContext*, este elemento não foi considerado por uma grande parte dos especialistas ou serviu propósitos diferentes dos aclarados, tendo sido antes usado para identificar a fase ou tarefas do processo construtivo em que determinado conceito se inseriria.

Assim, apesar de este elemento nos parecer relevante e útil para o bom entendimento do uso dos termos e para a clarificação das propostas, não o foi do ponto de vista dos especialistas, cujos contributos, como pudemos observar, se concentraram a outros níveis. A identificação de contextos de uso não foi, assim, entendida como uma tarefa prioritária para a conceptualização e para a localização.

Quando questionados sobre o uso deste elemento, foi identificado como acessório, em resultado da complexidade do processo e do tempo exigido para o seu desenvolvimento. AA_{es} sugeriu, como referimos antes, que este elemento fosse preenchido com informação obtida pelo terminólogo, o que não ia de encontro aos propósitos definidos para *usageContext*, cuja intenção era a de permitir aos especialistas documentar as suas decisões.

Dado que as limitações de tempo são características transversais aos processos de conceptualização e de localização, decidimos introduzir uma pequena modificação ao ambiente de conceptualização e localização e considerar, para utilizações futuras, este elemento como opcional. A sua inclusão deve, assim, ser ponderada de acordo com os diferentes contextos de conceptualização, as metodologias definidas, os actores e limites temporais.

Consideramos, no entanto, que a inclusão dos contextos de uso, obtida, por exemplo, a partir da sua extracção com recurso a um *corpus* textual, pode contribuir para apoiar a conceptualização do domínio, como referimos no ponto 1.2.2, pelo que, procuraremos, em desenvolvimentos futuros do método, disponibilizar este tipo de funcionalidade.

CONCLUSÃO

Conclusão

O presente trabalho de investigação visou propor um método para apoiar a especificação multilingue de ontologias de domínio, através do recurso à localização. Este método foi desenvolvido de modo a ser aplicado durante o processo de conceptualização de uma ontologia, após um processo de conceptualização inicial, sendo o desenvolvimento daquele enquadrado no contexto de uma rede colaborativa multilingue.

A construção do método, assente numa metodologia de base terminológica, seguiu uma perspectiva de análise interlinguística como base para a (re)conceptualização e localização do domínio, num enquadramento teórico que privilegia a abordagem conceptual à representação do conhecimento, mas que considera os recursos e contributos resultantes da abordagem linguística.

O método proposto definiu e conjugou etapas e elementos com o objectivo de (1) apoiar a especificação multilingue de ontologias, tendo por base a localização da representação do conhecimento disponibilizada em mapas conceptuais; (2) estruturar o processo de localização, através da identificação de um fluxo de trabalho e de um conjunto de ferramentas que assistissem os especialistas e o terminólogo e, eventualmente, outros actores participantes no processo; (3) criar melhores condições para fomentar a interacção entre os interlocutores envolvidos na conceptualização; (4) identificar caminhos que auxiliassem a integração de informação multilingue durante o processo de conceptualização, no contexto de uma rede colaborativa e de um ambiente multilingue de representação do conhecimento.

Para atingir estes objectivos, começámos por observar as lacunas patentes no acesso à informação e à comunicação multilingue; lacunas inerentes à dificuldade de utilização das diferentes línguas para a criação, gestão e disponibilização de conteúdos multilingues, tendo reflectido sobre o modo como este cenário pode representar um problema para as organizações e redes internacionais, bem como sobre a necessidade de as suprir de modo a fazer face ao crescente desenvolvimento de ambientes interorganizacionais, caracterizados pela multiculturalidade e pelo multilinguismo.

Tendo reconhecido o esforço significativo desenvolvido para minorar estas lacunas por diferentes instituições e em diferentes projectos, identificámos, no entanto, um conjunto de barreiras linguísticas que permanece ainda sem resolução, sobretudo quando o nosso enfoque recai sobre a comunicação multilingue em domínios específicos do saber ou sobre a representação multilingue do conhecimento.

Reconhecemos, neste contexto, a cada vez maior relevância assumida pelas ontologias como recurso privilegiado para a representação do conhecimento de um domínio específico. Tendo assumido que as ontologias são independentes de qualquer língua, defendemos que a criação e o acesso ao conhecimento continuarão a basear-se na língua, pelo que as ontologias deverão possuir, conseqüentemente, um carácter multilingue. Tal conduz à necessidade crescente da existência de recursos linguísticos, terminológicos e ontológicos multilingues, bem como de ferramentas e ambientes de localização que contribuam para ultrapassar as dificuldades de comunicação e de disseminação da informação e do conhecimento.

Observámos, por outro lado, que o estabelecimento de espaços multilingues de partilha de informação - como as redes sociais e as redes colaborativas, que se caracterizam pela heterogeneidade e multiculturalidade – cria a necessidade da existência de novos tipos de sistemas de organização do conhecimento que facilitem e tornem mais eficiente o acesso aos recursos de informação existentes e a sua reutilização.

Entendemos, no entanto, que ainda se verifica uma falha de comunicação entre as representações formalizadas do conhecimento, como as ontologias, desenvolvidas com o intuito de facilitar a reutilização deste conhecimento e a interoperabilidade na Web, e os utilizadores dos sistemas de informação que recorrem à sua língua (natural) para aceder ao conhecimento.

Apontámos, como forma de ultrapassar esta falha, a necessidade de explorar e desenvolver mais aprofundadamente as técnicas de localização em adequação a estas necessidades e a este contexto, bem como em conjugação com a promoção de novas abordagens, de carácter multidisciplinar, com recurso a modelos de conhecimento do mundo inspirados, por exemplo, nas redes sociais, visão que procurámos desenvolver ao longo deste estudo.

Atendendo a este contexto, apontámos também a necessidade de representação sistemática e dinâmica da informação e do conhecimento especializado, que entendemos ser relevante para todos os interlocutores de uma comunidade, dada a necessidade de acesso ao conhecimento de modo coerente entre os diferentes agentes e a cada vez maior urgência no uso de dados estruturados e normalizados em sistemas de informação e comunicação, que permitam transformar a informação existente no seio de uma organização em conhecimento acessível, disseminável e reutilizável. Destacámos, neste processo, o papel fundamental que a língua e a terminologia assumem na transmissão do conhecimento, bem como o contributo que oferecem à sua eficiente conceptualização e localização.

Tendo por base estas perspectivas e partindo de um ponto de vista das ciências da linguagem, debruçámo-nos, em seguida, sobre o papel desempenhado pela terminologia na representação do conhecimento, mais especificamente, no desenvolvimento do processo de conceptualização e da construção de ontologias representadas em mais do que uma língua. Analisámos, para tal, o processo de conceptualização, que vemos como central ao processo de construção de uma ontologia, e a interligação entre o uso de práticas e metodologias de base terminológica e o eficiente desenvolvimento daquele processo, sobretudo numa fase de especificação semiformal da conceptualização.

Tendo definido o que entendemos por conceptualização e por ontologias, descrevemos, em seguida, a relação da terminologia, das suas práticas e dos seus métodos, com o desenvolvimento de ambas. Partindo de uma análise dos pontos que aproximam e que distanciam terminologias e ontologias em termos de objectivos e metodologias, salientámos a importância do recurso à terminologia e às abordagens de base terminológica como ponto de partida para a análise e conceptualização do domínio e para a construção de ontologias.

A terminologia é uma disciplina relacionada com o estudo do conhecimento, cujo objecto de estudo são os conceitos e os termos que os designam. Partindo deste prisma, e dado que uma conceptualização é desenvolvida a partir da análise dos conceitos, acreditamos que tanto os especialistas de domínio envolvidos no processo colaborativo como os terminólogos focalizam a sua atenção no mesmo objecto – o conceito -, ainda que partindo de pontos de vista diversos. Esta dicotomia de análises e propósitos implica, assim,

que os métodos terminológicos sejam utilizados em adequação ao ambiente comunicativo específico e às aplicações que o enquadram, de modo a apoiar a fase de conceptualização de ontologias.

O contributo da terminologia para o processo de construção de ontologias tem seguido duas perspectivas teóricas e metodológicas distintas: a perspectiva conceptual, de base onomasiológica, baseada numa abordagem extralinguística, e a perspectiva linguística, de base semasiológica, em que prevalece o ponto de vista linguístico. Atendendo a esta diferença de partida na análise e organização do conhecimento especializado, reflectimos sobre as duas perspectivas teóricas, de modo a aclarar as abordagens propostas por cada.

Tal como afirmámos, e dadas as limitações apontadas à utilização de uma perspectiva linguística para a construção de uma representação do conhecimento de um domínio, é nossa opinião que o recurso a uma perspectiva conceptual é a mais adequada à implementação do método para a especificação multilingue de ontologias, devendo a representação do conhecimento e a sua localização partir daquela perspectiva.

Como pudemos observar, a primazia da utilização da abordagem conceptual influenciou positivamente o desenvolvimento dos processos de conceptualização e localização, tendo tornado o recurso à abordagem linguística uma necessidade complementar e com objectivos muito específicos.

De facto, e como resulta do trabalho realizado e da experimentação do método, concluímos que embora possam ser complementares, a ordem de utilização da perspectiva conceptual e da perspectiva linguística não pode ser arbitrária, devendo o processo de conceptualização partir da primeira e apoiar-se no especialista, tal como havíamos proposto.

No entanto, e na nossa opinião, as duas perspectivas não ocupam planos antagónicos ou mutuamente exclusivos. Consideramos mesmo que uma abordagem mista, que resulte do uso complementar das duas perspectivas, de modo flexível e em adequação às necessidades e propósitos de análise do conhecimento de um domínio específico, e que tenha em conta o contexto comunicativo específico em que se desenrolam a conceptualização e a localização, pode conduzir a uma análise mais aprofundada e a resultados mais profícuos na representação desse mesmo conhecimento.

Entendemos, por outro lado, que a metodologia mista utilizada neste estudo salientou resultados promissores para a conceptualização do domínio e para a especificação multilingue da ontologia, apontando para a necessidade de considerar o *corpus* textual como complemento ao desenvolvimento de uma abordagem conceptual, dadas as possibilidades oferecida pelo texto e pelos restantes recursos disponíveis para a extracção e o confronto de informação sobre o domínio, por um lado, e para a organização do conhecimento, por outro.

Como decorre da análise dos resultados da aplicação do método, podemos concluir que, no caso da especificação multilingue de uma ontologia, o recurso às duas perspectivas é uma mais valia no desenvolvimento de uma abordagem interlinguística ainda que a perspectiva conceptual, deva, também aqui, assumir a primazia. De facto, numa abordagem interlinguística o processo de escolha dos termos não pode partir das formas, mas dos conceitos, não podendo, conseqüentemente, ser uma tradução ou uma localização de nomes, mas, antes, um processo de selecção das denominações naturais que em cada língua correspondam a um conceito de especialidade, tal como expresso pelos especialistas em situações de comunicação profissional.

Defendemos, por isso, por contraste com o trabalho de organização do conhecimento centrado exclusivamente na visão do especialista, um trabalho interdisciplinar, que inclua a presença da terminologia e do terminólogo na fase semiformal de especificação de uma conceptualização e na sua posterior localização, e que conjugue, complementarmente, as perspectivas conceptual e linguística, de modo a conduzir a uma melhor organização do conhecimento.

Considerámos, ainda, que o recurso às metodologias e práticas terminológicas no âmbito da conceptualização contribui para apoiar o processo comunicativo que se desenrola entre os especialistas ao longo do desenvolvimento do trabalho colaborativo, e entre estes e os restantes interlocutores do processo, promovendo, conseqüentemente, uma melhor gestão da dinâmica gerada em torno da elicitação dos conceitos e da sua representação.

Um elemento central a este estudo é a localização. Da análise ao conceito e ao processo de localização, concluímos que a terminologia, a tradução e a localização mantêm, enquanto disciplinas, uma relação de proximidade que se pode observar em diversos aspectos. Tanto a terminologia como a localização são campos interdisciplinares e possuem,

na sua base, elementos dos campos das ciências cognitivas, das ciências da linguagem e das ciências da comunicação.

A presença coincidente destes enfoques epistemológicos resulta, como descrevemos, do facto de que o objecto das ciências dedicadas à comunicação e à informação se sintetizar a partir da descrição e explicação de três elementos: as categorias do conhecimento; as unidades que permitem designar estas categorias; e o conhecimento que é expresso e transmitido por estas unidades. Da conjugação destes elementos resulta, na nossa perspectiva, uma relação simbiótica entre o estudo da terminologia e o estudo da localização, que coloca a terminologia no centro do processo de localização e, por extensão, do processo de localização de ontologias.

Quando nos debruçamos sobre a localização como processo, percebemos, assim, a necessidade de ter presente na sua análise e no seu estudo os contributos da terminologia. Se na análise do processo de localização nos deparamos com a questão sobre a forma como o conhecimento especializado é disponibilizado e transmitido, a resposta surge, como apontámos, considerando a terminologia como o modo privilegiado para essa transmissão, graças à profundidade de análise da língua de especialidade oferecida pelas suas práticas e metodologias.

O recurso à terminologia sustenta, como observámos, o eficiente desenvolvimento do processo de localização e da localização de ontologias e contribui para a garantia da qualidade e consistência dos conteúdos localizados, constituindo a terminologia o elemento de base de transferência de conhecimento entre o produtor e o utilizador final, os quais, no caso da localização de uma ontologia, se podem entrecruzar, como aconteceu no nosso caso de estudo.

A terminologia desempenha, como procurámos demonstrar, um papel decisivo no processo de localização, uma vez que está relacionada com as designações atribuídas aos diferentes elementos da ontologia e é o veículo que permite que os termos da ontologia possam ser expressos em mais do que uma língua. Como ficou claro, no contexto de trabalho bilingue e multilingue em que se insere a localização de ontologias, a terminologia desempenha um papel importante ao assistir os terminólogos e os especialistas nas tarefas de acesso ao conhecimento, na criação do sistema de conceitos, na formulação de

definições e, finalmente, na procura e no estabelecimento de equivalentes entre os termos das diferentes línguas.

A localização não é, como pretendemos demonstrar, apenas uma questão de língua. É, antes, um processo que envolve um conjunto de actividades e actores com a preocupação de possibilitar a comunicação entre diferentes línguas a partir da correcta gestão e transferência de informação e conhecimento, efectuadas com base na análise sistemática dos conceitos de um domínio de especialidade, nomeadamente, enquanto processo integrante da gestão do conhecimento baseado em ontologias.

Fruto, cremos, da complexidade inerente ao processo de localização de ontologias, só mais recentemente se notou uma evolução nos sistemas técnicos e abordagens teóricas disponíveis para desenvolver este processo, bem como no desenvolvimento de alguns princípios epistemológicos que possibilitaram a abertura de novos campos de reflexão e análise nesta área.

Esta complexidade tem resultado na escassez de ontologias formalizadas que contenham, associada, informação linguística em mais do que uma língua natural; escassez que provém de vários factores, a começar pela dificuldade inerente à escolha das metodologias a adoptar para a conceptualização e representação do conhecimento e pela existência em número diminuto de ferramentas específicas que apoiem este processo. Outras dificuldades, mais específicas, advêm sobretudo de problemas de ordem linguística.

Acresce a estas que o processo de localização de uma ontologia, artefacto específico para a representação do conhecimento de um domínio e construído num contexto e com um propósito particular, tem que atender, na nossa perspectiva, a outros requisitos. Destes destacam-se os relacionados com a definição e delimitação do domínio ou subdomínio(s) a conceptualizar; a tipologia, classificação e conversão de recursos semânticos a adaptar para tornar o processo célere e enriquecer a ontologia; a existência de metodologias que apoiem a integração entre os processos de localização e conceptualização; a limitação temporal normalmente imposta ao desenrolar destes processos; e o modo de integração e (re)utilização de recursos e ferramentas linguísticas já disponíveis.

Nesse sentido, desenvolvemos uma análise ao processo de localização de ontologias, em que explicitámos as diferentes dimensões a ter em conta no seu desenvolvimento e as

dificuldades que lhe são inerentes e que se estendem à gestão dos elementos multilingues de uma ontologia. Identificámos, nesta análise, as diferentes metodologias disponíveis para apoiar a representação de diferentes línguas numa ontologia, bem como alguns métodos e ferramentas construídos para apoiar a localização de ontologias.

De modo a contribuir para uma melhor descrição da relação entre o processo de localização e o de localização de ontologias, estabelecemos, ainda, um conjunto de analogias, que identificámos a vários níveis, de modo a melhor contextualizar o desenrolar do processo de localização de ontologias e para perspectivar os elementos comuns que aproximam este processo do processo tradicional de localização.

Com base na análise efectuada, desenvolvemos a perspectiva de que a localização de ontologias é um processo de adaptação linguística e cultural dos conteúdos de uma ontologia aos requisitos de uma língua e de uma cultura alvo, elaborado de modo a corresponder às suas necessidades específicas de acesso e gestão do conhecimento.

O recurso às metodologias de base terminológica surge como um elemento essencial, no nosso ponto de vista, para a resolução de um conjunto questões que se colocam no desenrolar do processo de especificação multilingue de uma ontologia. Por outro lado, a construção de estruturas conceptuais representativas do conhecimento do domínio é, há muito, prática corrente no trabalho e na pesquisa terminológica, pelo que, os modelos terminológicos referentes à representação do conhecimento do domínio são uteis para explorar a possibilidade de transformar aquelas estruturas em ontologias.

Como destacámos ao longo deste estudo, é nossa opinião que no desenvolvimento de ontologias se tem dado uma importância crescente à fase da construção partilhada de representações da conceptualização do domínio, desenvolvida por grupos de especialistas, fase que não mereceu ainda a devida atenção no que se refere à relação entre a construção partilhada de uma ontologia e o seu enriquecimento com elementos multilingues, sobretudo quando contextualizada no seio de uma rede colaborativa multilingue.

Neste contexto, a terminologia e as suas ligações à representação do conhecimento sugerem, como afirmámos, caminhos alternativos para o tratamento da problemática da criação e manutenção de ontologias em várias línguas, tendo por base a conjugação de contributos dos dois tipos de trabalhos – o terminológico e o da representação do

conhecimento. Estes partilham o facto de servirem propósitos idênticos, de que destacamos aqui o de procurarem fornecer uma representação partilhada do conhecimento sobre uma parte específica do mundo a um conjunto de diferentes utilizadores, representação que procura ser independente de uma qualquer língua natural.

Tendo em conta este contexto, os princípios e pontos de vista acima enunciados, bem como as lacunas identificadas ao longo da análise do estado da arte, desenvolvemos a ideia de construir um método de apoio à especificação multilingue de ontologias na sua fase de conceptualização, com recurso à localização. Para a concepção deste método contribuiu de forma importante a análise dos pressupostos teóricos, expostos e debatidos ao longo do presente trabalho, bem como a observação directa do funcionamento de uma rede colaborativa.

Este método foi desenvolvido com o fim de, por um lado, contribuir para a resolução de parte dos problemas e lacunas identificados – de que se destacam a dificuldade inerente à escolha de metodologias a adoptar para a representação do conhecimento num ambiente de conceptualização e a localização de ontologias e o número diminuto de ferramentas e ambientes específicos que apoiem este processo - e, por outro, auxiliar aos processos de construção partilhada de uma ontologia e da sua localização para diferentes línguas, tendo em conta a participação dos especialistas de domínio.

A concepção do método considerou também o contexto particular no qual seria aplicado - uma rede colaborativa, espaço multilingue singular no qual cooperam vários interlocutores num esforço comum de estruturação e representação do conhecimento relativo a um domínio. Com efeito, pretendeu-se proporcionar um espaço de trabalho que, através da inclusão e integração dos meios potenciase as dimensões expressivas/comunicativas mais complexas, projectando um espaço de autonomia e de partilha de conhecimento, construído a partir de uma abordagem necessariamente transdisciplinar que considerou os aspectos do acesso e a partilha de conhecimento, da sua representação, localização e da sua posterior especificação formal.

Com o fim de implementar e validar o método proposto para a especificação multilingue de ontologias, desenvolvemos uma experiência realizada com base num caso de estudo. Esta experiência foi realizada no âmbito do projecto europeu H-Know e consistiu na aplicação do método à especificação multilingue de uma ontologia do domínio da

reabilitação; ontologia construída no âmbito da rede colaborativa criada no contexto do projecto.

Esta experiência, cujos resultados foram alvo de análise prévia no ponto 4.2.5.2, permitiu também verificar a validade das nossas hipóteses de partida. De facto, o conjunto das acções desenvolvidas ao longo das diferentes fases da experiência permitem-nos validar a primeira hipótese que colocamos no início deste estudo, nomeadamente a de que a localização, ao ser realizada numa fase subsequente à da conceptualização e anterior à da especificação formal de uma ontologia, poderia representar uma ajuda à construção da conceptualização em mais do que uma língua e enriquecer o processo de partilha do conhecimento durante a construção de uma ontologia.

Como demonstrámos, o uso de um processo consecutivo de conceptualização e localização pode concorrer para o acesso mais imediato a um conjunto mais vasto de conhecimento sobre o domínio e permitir aos especialistas uma negociação e estruturação do conhecimento do domínio potencialmente mais completa e, em consequência, uma especificação mais granular da(s) ontologia(s).

Tal acontece quando se proporcionam as condições para que se desenvolvam representações conceptuais do conhecimento utilizando, para a sua descrição, as línguas de trabalho dos diferentes especialistas da rede. Por outro, criam-se condições para a obtenção de um *feedback* mais eficaz sobre a conceptualização do domínio a partir das representações semiformais de conhecimento, admitindo a sua eventual reconceptualização, ao levar a que se activem e considerem as características específicas das línguas e culturas em presença, e, finalmente, ao promover uma participação mais activa dos especialistas na construção e na localização da ontologia.

O desenvolvimento iterativo e consecutivo destes processos constituiu-se, assim, como um factor de intercompreensão capaz de auxiliar no processo de negociação, ao diminuir o grau de ambiguidade e ao contribuir para a eliminação de elementos que possam concorrer para a confusão conceptual e terminológica. Como aventámos, o processo de localização permitiu, neste contexto, que se considerassem não só os elementos que constituem o sistema conceptual – conceitos e relações - como acontece quando a localização ocorre a partir de uma ontologia formal, mas antes, e sobretudo, que os actores envolvidos no processo tivessem em conta a representação conceptual do conhecimento, a

partir da sua visualização sob a forma de mapas conceptuais, o que conduziu a uma melhor percepção e análise da interligação entre os conceitos e do seu valor.

Verificámos, assim, tal como sugerimos na segunda hipótese²¹⁴ que propusemos, que a localização representa, enquanto processo, um elemento que contribui para a construção mais efectiva de uma conceptualização do conhecimento do domínio, podendo facilitar, posteriormente, os processos de especificação formal e de manutenção da ontologia.

De facto, embora o especialista parta de uma perspectiva conceptual, baseando-se esta fundamentalmente em dados extralinguísticos, sente sempre a necessidade de mediação da língua natural na especificação informal da conceptualização e na localização. Como observámos, o desenvolvimento do processo de localização levou a que, por via do discurso, os especialistas explicitassem de forma mais precisa o conhecimento, através, por exemplo, do modo como reflectiram sobre a conceptualização inicial ou da forma como colocaram as dúvidas ao longo do processo de escolha dos equivalentes.

Como pudemos observar também, os processos de (re)conceptualização e de localização podem envolver um grande número de actores da rede colaborativa. Esse número pode implicar, na cadeia contínua de contributos e de negociação que se gera, alguma perda de percepção do significado original de um conceito. Esta possibilidade reforça a necessidade da presença do terminólogo, pelo auxílio que pode prestar neste processo de acesso, transferência e organização do conhecimento entre os actores da rede colaborativa ao mediar entre as diversas perspectivas e propostas.

A necessidade da presença de um terminólogo é referida na opinião expressa por alguns dos especialistas durante a entrevista que conduzimos, como JPM_{pt} ao afirmar: (...) *Neste processo o linguista tem um papel de mediação – suscita a discussão, desestabiliza, retira o especialista da sua zona de conforto, é provocador, promove o ambiente de negociação. Permite, com a sua presença, uma maior reflexão sobre a representação.*

A existência, tal como proposto no método, de um intermediário pode, por outro lado, minorar alguma eventual incompreensão que exista numa análise inicial da representação conceptual e apoiar o desenvolvimento do processo, uma vez que este pode desempenhar um papel activo no esclarecimento de dúvidas, na identificação de fontes ou

²¹⁴ A validade desta hipótese foi, por outro lado, verificada a partir das respostas obtidas com base na realização da entrevista semiestruturada aos especialistas, como descrevemos no ponto 4.2.5.3.

na explicitação do raciocínio que esteve na base de cada mapa, conduzindo, em princípio, a uma mais rápida obtenção de consenso.

Do mesmo modo, a disponibilização de elementos de base terminológica, como a definição dos conceitos em língua natural e a identificação das suas fontes, surge como um contributo importante para a existência de uma referência comum e transversal, facilmente acessível e que pode contribuir para um desenvolvimento mais célere do processo ao evitar alguma dispersão.

Constituindo as ontologias um meio para operacionalizar modelos conceptuais que representam o conhecimento do domínio, torna-se, na nossa opinião, necessário utilizar e integrar, no seu desenvolvimento, métodos de representação conceptual focalizados no conhecimento do domínio, destacando-se neste processo, como demonstrámos, as abordagens propostas pela terminologia à conceptualização do domínio, uma vez que se dedicam, com grande especificidade, à análise da natureza do conhecimento do domínio, i.e. da sua estrutura conceptual.

A análise do trabalho levado a cabo pelos especialistas, os dados obtidos a partir das propostas e sugestões que apresentaram e das diversas interações que se tornaram possíveis ao longo da experimentação deste método, foram de encontro à terceira hipótese que colocámos, quando consideramos que o processo de conceptualização de uma ontologia de domínio no seio de uma rede colaborativa multilingue condicionaria positivamente o processo de localização, tal como é percebido tradicionalmente.

Esta hipótese foi colocada dado que, como era nossa percepção, não estaríamos perante um processo tradicional de localização de termos isolados ou descontextualizados, mas antes em presença de uma representação conceptual inicial que se constituiria como o elemento nuclear ao processo de localização de uma ontologia, elemento sobre o qual os especialistas agiriam e que sustentaria as suas propostas de equivalentes e de reconceptualização.

De facto, o processo de localização centrou-se não só nos elementos individuais que constituíam cada sistema conceptual – conceitos, designados por termos, e relações que os interligam – e na procura dos seus equivalentes, mas também na (re)análise da representação conceptual. Esta permitiu a comparação e a reestruturação das

representações conceptuais a partir da análise interlinguística dos conceitos e dos termos que os designam, seguindo, assim, os princípios de uma abordagem conceptual, tendo o especialista agido de forma a contribuir para uma representação conceptual que expressasse o conhecimento do domínio sem deixar de ter em conta também a sua comunidade, à qual se destinava a ontologia que analisou, (re)conceptualizou e localizou.

A realização da experiência, no âmbito do caso de estudo referido, permitiu, finalmente, avaliar o método, identificar as suas limitações e perceber as melhorias a introduzir no seu fluxo de trabalho e nos seus componentes, de modo a melhor apoiar o processo e o fluxo de acesso, partilha e gestão da informação e do conhecimento expresso e representado em mais do que uma língua, bem como a capacidade de comunicação entre os diferentes interlocutores da comunidade.

As opiniões expressas pelos especialistas permitiram, finalmente, concluir que os pressupostos e objectivos com que construímos o método, reflectidos no seu uso por esses mesmos especialistas, se adequam ao contexto do desenvolvimento de um processo de localização na fase de conceptualização de uma ontologia de domínio no contexto de uma rede colaborativa multilingue, tendo, no caso específico do nosso estudo, o recurso ao método conduzido à obtenção de resultados relevantes para a construção e especificação multilingue da ontologia.

A análise levada a cabo neste estudo, ainda que tendo tido como ponto de partida a especificação multilingue de uma ontologia, desenvolvida no contexto de um projecto e de um domínio particular, procurou adoptar uma postura de transpor para o caso de estudo pressupostos, abordagens e pontos de vista gerais passíveis de serem aplicados em casos futuros.

Os resultados obtidos, que consideramos positivos e satisfatórios, não implicam, contudo, a não identificação ou a não existência de limitações no método proposto, anteriormente descritas, como a existência de elementos cujo uso foi diminuto, e de outras que ultrapassam, necessariamente, o âmbito do presente trabalho de investigação e que obrigariam a um esforço considerável para o seu desenvolvimento e a sua execução, bem como ao desenvolvimento de competências que não possuímos, como as de programação.

A identificação dessas lacunas, o trabalho de investigação realizado e o desenvolvimento da experiência descrita no capítulo anterior, permitem sugerir algumas orientações de trabalho futuro que nos parecem pertinentes e que poderão contribuir para a continuidade da investigação nesta área.

Assim, procuraremos, no contexto da especificação multilingue de ontologias e do uso de um espaço de conceptualização e localização baseado em mapas conceptuais, avaliar as possibilidades de introdução de uma ferramenta que possibilite o acesso aos contextos de uso a partir quer de um *corpus* monolingue quer de um *corpus* comparável e que permita aos especialistas o acesso em tempo real àqueles contextos, de modo a apoiar as diferentes actividades inerentes ao processo de (re)conceptualização e de localização.

Procuraremos, também, avaliar a possibilidade de disponibilização, no espaço de trabalho proposto, de um serviço que permita o acesso directo e simultâneo a um conjunto de aplicações de tradução automática e de bases de dados terminológicas e lexicais, de modo a permitir ao especialista um acesso mais imediato aos resultados da consulta, tornando o processo de localização mais ágil e o tipo de resultados obtidos potencialmente mais diversificado e mais rico.

Este serviço deve ser desenvolvido de modo a considerar a hipótese de localização de um termo singular ou de localização dos termos da representação conceptual na sua totalidade e fornecer ao especialista a possibilidade de escolher e validar o equivalente mais adequado ou de ignorar os resultados obtidos. Poderá, finalmente, disponibilizar ao especialista um meio para seleccionar mais do que um dos equivalentes obtidos e classificá-los, por exemplo, em função de critérios de equivalência previamente definidos.

Por outro lado, ainda no mesmo contexto, procuraremos avaliar as vantagens de incorporar no espaço de conceptualização e localização mecanismos formais de apoio à negociação e à decisão, tal como sugerido por alguns dos especialistas. Estes mecanismos poderiam auxiliar, por exemplo, no processo de escolha e validação dos equivalentes a validar, quando dele participassem mais do que um especialista com a mesma língua de trabalho.

As plataformas dirigidas às redes interorganizacionais procuram a promoção da colaboração e da difusão do conhecimento na rede, num ambiente que deve transcender as

barreiras geográficas e linguísticas. Para tal devem, a nosso ver conjugar serviços que combinem a gestão de conteúdos digitais com tecnologias de localização e de pós-edição, que prevejam a participação dos especialistas de domínio, independentemente da sua formação, e fomentem o acesso e disponibilização de conteúdo em várias línguas. Nesse sentido, procuraremos, finalmente, estender a aplicação do método proposto a outros contextos e redes colaborativas ou comunidades de prática, de modo a avaliar a sua aplicabilidade em cenários distintos do que suportou a nossa análise e experiência e a desenvolver o método a partir dos resultados obtidos pela sua aplicação nesses contextos.

Pretendemos, ainda, com esta aplicação do método em novos cenários, no âmbito de contextos multilingues e transculturais que envolvem o uso de comunicação especializada, observar se se produzem alterações na natureza das práticas terminológicas e, em especial, no trabalho desenvolvido pelos terminólogos, tendo em conta os novos desafios que se lhes colocam.

A implementação das ferramentas e serviços descritos e o seu teste, em contextos de trabalho multilingues, é uma tarefa complexa e extensa que só será exequível, a nosso ver, a médio ou longo prazo e que terá que considerar os avanços teóricos e tecnológicos entretanto produzidos no âmbito de projectos correntemente em desenvolvimento, quer na área da terminologia, quer na área da representação do conhecimento, bem como as possíveis interligações que se forjem ou se reforcem entre as duas áreas, cuja proximidade e trabalho conjunto consideramos essencial à efectiva e eficiente organização, disseminação e acesso ao conhecimento especializado.

BIBLIOGRAFIA

Bibliografia

- Aguado de Cea, G., Montiel-Ponsoda, E., e Ramos Gargantilla, J.A. (2007). *Multilingualidad en una aplicación basada en el conocimiento*. Em TIMM, monográfico para la revista SEPLN, nº38, pp. 77-97.
- Antia, B. (2000). *Terminology and language planning: an alternative framework of practice and discourse*. John Benjamins Publishing Company, Amsterdam/Philadelphia.
- Antia, B., Budin, G., Picht, H., Rogers, M., Schmitz, K., e Wright, S. (2005). Shaping Translation: A View from Terminology Research. In *Meta: journal des traducteurs* Vol. 50, nº 4, 2005. Acedido em 29 de outubro de 2008 em: <http://id.erudit.org/iderudit/019907ar>
- Arntz, R. e Picht, H. (1995). *Introducción a la Terminología*. Biblioteca del Libro, Fundación Germán Sánchez. Ediciones Pirámide S.A., Ruipérez, Madrid. p. 190.
- Aubenque, P. (1962). *Le problème de l'Être chez Aristote*. PUF, Paris.
- Aussenac-Gilles, N., Biebow, B., e Szulman, S. (2000). *Revisiting Ontology Design: A Methodology Based on Corpus Analysis*. EKAW 2000. pp 172-188.
- Babych, B., Hartley, A., e Sharoff, S. (2007). *Translating from under-resourced languages: comparing direct transfer against pivot translation*. Em Proceedings of MT Summit XI, Copenhagen. pp. 10-14.
- Barrasa, J. (2007). *Modelo para la definición automática de correspondências semânticas entre ontologías y modelos relacionales*. Unpublished doctoral dissertation, Facultad de Informatica, Universidad Politécnica de Madrid. Madrid, Spain. Acedido em 15 de agosto de 2009 em: <http://eprints.eemcs.utwente.nl/7146/>
- Barros, S. (2007). *A relação genérica num corpus de direito constitucional português: identificação de contextos para a delimitação de candidatos a termo*. FCT - UNL. Universidade Nova de Lisboa, Lisboa.
- Barros, S., Costa, R., Soares, A., e Silva, M. (2012). *The role of terminological principles and methods in a framework for a collaborative development of semi-formal ontologies*. Soares, A., Costa, R. e Pereira, C. (eds). Em Proceedings of the Workshop on

Collaboration in Terminology and Knowledge Application. LREC, 2012. Istambul. ELRA. Acedido em 15 de junho de 2012 em <http://www.lrec-conf.org/proceedings/lrec2012/workshops/08.LREC%202012%20Terminology%20Proceedings.pdf>

Bateman, J. (2004). The place of language within a foundational ontology. Em Achille C., Varzi e Vieu. L. (eds.), *Formal Ontology in Information Systems: Proceedings of the Third International Conference (FOIS 2004)*, pp. 222–233. IOS Press, Amesterdam.

Bassnett, S. e Trivedi, H. (1999). *Post-colonial Translation – Theory and Practice*. Routledge, London/New York.

Beaugrande, R. (1987). Determinacy distributions in complex systems: science, linguistics, language, life. Em *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 40. pp. 145-188.

Beaugrande, R. e Dressler, W. (1994). *Introduction to Text Linguistics*. Longman, London.

Béjoint, H. (1997). Modele Relationnel, Definition et Denomination. Em *Autour de la Dénomination*. Presses Universitaires de Lyon, Lyon.

Béjoint, H. e Thoiron, P. (2000). *Le Sens en Terminologie*. Presses Universitaires de Lyon, Lyon.

Benjamins, V., Contreras, J., Corcho, O., e Gómez-Pérez, A. (2002). Six Challenges for the Semantic Web. In Proceedings of KR 2002, Knowledge Representation Semantic Web Workshop. Acedido em 22 de Junho de 2010 em: http://www.dia.fi.upm.es/~ocorcho/documents/KRR2002WS_BenjaminsEtAl.pdf

Bergamaschi, S. (2003). SEWASIE: Semantic Webs and Agents in Integrated Economies. WP2 –Semantic Enrichment and Integration. In Deliverable D2.1: *Specification of the general framework for the multilingual semantic enrichment processes and of the semantically enriched data stores*. Acedido em 14 de Setembro de 2009 em: <http://www.sewasie.org/documents/D-2-1-semantic-enrichment.pdf>

Bergenholtz, H. e Tarp, S. (1995). *Manual of Specialised Lexicography. The Preparation of Specialised Dictionaries*. John Benjamins Publishing Company, Amsterdam/Philadelphia.

- Berners-Lee, T., Hendler, J., and Lassila, O. (2001). *The Semantic Web*. Scientific American.
- Bessé, B. (1997). Terminological Definitions. Em *Handbook of Terminology Management – Volume 1: Basic Aspects of Terminology Management*. John Benjamins, Amsterdam/Philadelphia.
- Biber, D. (1993). Representativeness in Corpus Design. Em *Literary and Linguistic Computing* 8. Acedido em 4 de Dezembro de 2010 em: <http://staff.um.edu.mt/albert.gatt/teaching/dl/biber93.pdf>
- Borst, W. (1997). *Construction of Engineering Ontologies*. PHD thesis, Institute for Telematica and Information Technology, University of Twente, Enschede, The Netherlands.
- Boteram, F. e Hubrich, J. (2008). *Towards a comprehensive international Knowledge Organization System Networked Knowledge Organization Systems NKOS*. ECDL-Conference, Århus.
- Bourigault, D., Jacquemin, C., e L’Homme, M. (Eds.). (2001). *Recent Advances in Computational Terminology*. John Benjamins Publishing Company, Amsterdam/Philadelphia.
- Bourigault, D., e Aussenac-Gilles, N. (2003). *Construction d'ontologies à partir de textes*. Paper presented at the TALN 2003 Batz-sur-Mer.
- Bowker, L. (1997). Multidimensional Classification of Concepts and Terms. Em *Handbook of Terminology Management – Volume 1: Basic Aspects of Terminology Management*. John Benjamins Publishing Company, Amsterdam/Philadelphia.
- Bowker, L., e Pearson, J. (2002). *Working with Specialized Language: A Practical Guide to Using Corpora*. Routledge, London.
- Braschler, M. Choukri, K., Ferro, N. Hanbury, A. Karlgren, J. Müller, H., Petras, V., Pianta, E., Rijke, M. e Santucci, G. (2010). A PROMISE for Experimental Evaluation. Multilingual and Multimodal Information Access Evaluation. Em International Conference of the Cross-Language Evaluation Forum, CLEF 2010. M. Agosti, N. Ferro, C. Peters, M. de Rijke and A. Smeaton (Eds.). Padua, Italy, September 2010, Springer. [Lecture Notes for Computer Science; 6360].

- Budin, G. (1994). Some hypotheses about concept representations. Em Proceedings of the 9th European Symposium on LSP, Bergen, Norway, 26 August, 1993. Bergen: Fagbokforlaget.
- Budin, G., Melby, Alan K. (2000). Aecessibility of Multilingual Terminological Resources - Current Problems and Prospects for the Future. Acedido em 24 de Abril de 2010 em: www.ttt.org/Salt/Athens_Paper.doc
- Budin, G. (2005). Ontology-driven translation management. In Knowledge Systems and Translation. Helle V. Dam, Jan Engberg, Heidrun Gerzymisch-Arbogast (Edt). Walter de Gruyter.
- Budin, G. (2007). Epistemological aspects of indeterminacy in postmodernist science. Em: B., Antia (eds). *Indeterminacy in Terminology and LSP*, John Benjamins Publishing Company. Amsterdam/Philadelphia. pp. 61-72. ISBN 978 90 272 2332.
- Budin, G., Alegre, C., Monfort, C. (2007a). Wide Information Network for Risk Management. Deliverable D2205.3. Risk management language & communication tools. Acedido em 24 de Abril de 2010 em: http://mgrm.univie.ac.at/fileadmin/user_upload/fak_translationswissenschaft/Projekte/mgrm/WIN-UMB-HLI-MULTH-PU-D2205.3_HLI_TOOLS_ONTOLOGY-V2.00_FR.pdf
- Budin, G. (2007b). *From Terminologies to Ontologies – Advances in Knowledge Organization*. TSS2007. Acedido em 15 de Junho de 2010 em: http://www.termnet.org/downloads/english/events/TSS2007_TerminologiesOntologies_more.pdf
- Buitelaar, P., Sintek, M., Kiesel, M. (2006). A Multilingual/Multimedia Lexicon Model for Ontologies. Em Proceedings of ESWC, Budva, Montenegro.
- Buitelar, P. *et al.* (2006a). LingInfo: Design and Applications of a Model for the Integration of Linguistic Information in Ontologies. In Proceedings of OntoLex 2006.
- Buitelaar, P., Declerck, T., Frank, A., Racioppa, S., Kiesel, M., Sintek, M., Engel, R., Romanelli, M., Sonntag, D., Loos, B., Micelli, V., Porzel, R., and Cimiano, P. (2006). Linginfo: Design and applications of a model for the integration of linguistic information in ontologies. In Proceedings of the OntoLex Workshop at LREC. ELRA.

- Buitelaar, P., Cimiano, P., Haase, P., & Sintek, M. (2009). Towards Linguistically *Grounded Ontologies*. In The 6th Annual European Semantic Web Conference (ESWC2009), Heraklion, Greece.
- Buitelaar, P., Cimiano, P., Montiel-Ponsoda, H., McCrae, J., and Declerck, T. (2011). Ontology Lexicalisation: The lemon Perspective. Em: Monique Slodzian, Mathieu Valette, Nathalie Aussenac-Gilles, Anne Condamines, Nathalie Hernandez, Bernard Rothenburger (eds.). *Workshop Proceedings of the 9th International Conference on Terminology and Artificial Intelligence*, pp 33-36, Paris, France, INALCO, INALCO, Paris.
- Cabré, M. T. (1999). Terminology – Theory, Methods and Applications. John Benjamins Publishing Company, Amsterdam /Philadelphia.
- Cabré, M. T. (1999a). Elementos para una teoría de la terminología: hacia un paradigma alternativo. In *La terminología, representación y comunicación: Elementos para una teoría de base comunicativa y otros artículos*. Barcelona:IULA, Universitat Pompeu Fabra, 1999b. pp. 69-92. ISBN 64-477-0673-7.
- Cabré, M. T. (1999b). Estandardització i interferència en terminologia. In *La terminología, representación y comunicación: Elementos para una teoría de base comunicativa y otros artículos*. Barcelona: IULA, Universitat Pompeu Fabra. pp. 39-54. ISBN 64-477-0673-7.
- Cabré, M. T., (2000). Sur la Représentation Mentale des Concepts: Bases pour une Tentative de Modélisation. in *Le Sens en Terminologie*. Presses Universitaires de Lyon, Lyon.
- Cabré, M. T. (2000). *Theories of Terminology*. Em Terminology 6 (1), pp 35–57.
- Cabré, M. T. (2003). Teorías de la terminología: de la prescripción a la descripción. Em *Innovazione lessicale e terminologie specialistiche*. Leo S. Olschki Editore, Florència. pp. 169-188.
- Cabré, M.T., Bach, C., Estopà, R., Feliu, J., Martínez, G., and Vivaldi, J. (2004). *The GENOMAKB project: towards the integration of concepts, terms, textual corpora and entities*. LREC 2004.
- Cabré, M.T. (2004b). La terminología en la traducción especializada. In Garcia, Consuelo e Yebra, Valentin (2004). *Manual de Documentación y termonologia para la traducción*

especializada. Arco/Libros, Madrid. Colección: Instrumenta Bibliológica. pp. 89-122. ISBN: 84-7635-578-5.

Cabré, M.T. (2005). La Terminología, una disciplina en evolución: pasado, presente y algunos elementos de futuro. In *Debate Terminológico* 1. París: RITERM (Red Iberoamericana de Terminología). Acedido em 2 de Janeiro de 2009 em: http://www.riterm.net/revista/n_1/index.htm

Cabré, M. T. (2007). Constituir un *corpus* de textos de especialidad: condiciones y posibilidades. Em Ballard, M., Pineira-Tresmontant, C. (eds.). *Les corpus en linguistique et en traductologie*. Artois Presses Université, Arras. pp 89-106. ISBN 978-2-84832-063-2

Calberg-Challot, M., Lerat, P., Roche, C. (2009). *Quelle place accorder aux corpus dans la construction d'une terminologie?*. In Actes TOTH. 33-52.

Calzolari N., e Soria C. (2005). A new paradigm for an Open Distributed Language Resource Infrastructure: the case of Computational Lexicons. Em *Proceedings of the AAAI Spring Symposium "Knowledge Collection from Volunteer Contributors (KCVC05)"*. Stanford, CA.

Camarinha-Matos, L., Afsarmanesh, H. (2005). Collaborative networks: a new scientific discipline. *Journal of Intelligent Manufacturing*, 16, pp 439–452. Springer Science, Business Media, Inc.

Cencioni, R., Rossi, K. (2008). Language based Interaction. Challenge 2 –Objective 2.2. DG Information Society and Media. ICT 2008. Lyon. Acedido em 12 de Julho de 2009 em: ftp://ftp.cordis.europa.eu/pub/fp7/ict/docs/language-technologies/events-20081126-cencioni-rossi_en.pdf

Ceusters, W. (2002). *Terminology and Ontology Management Systems*. Acedido em 17 de Março de 2010 em: Ingenieur.kahosl.be/werner%20%ceusters/swa2002

Ceusters, W. (2008). Solving Crimes using Referent Tracking: Crime Terminology. UB Discovery Seminar 035158/UE 141 MMM. Acedido em 12 de Março de 2010 em: <http://www.referent-tracking.com/RTU/>

- Cimiano, P., Haase, P., Herold, M., Mantel, M., and Buitelaar, P. (2007). LexOnto: A Model for Ontology Lexicons for Ontology-based NLP. In *Proceedings of the OntoLex07. Workshop held in conjunction with ISWC'07*.
- Cimiano, P., Montiel-Ponsoda, E., Buitelaar, P., Espinoza, M., and Gómez-Pérez, A. (2010). A note on ontology localization. *Journal of Applied Ontology*, 5(2), pp 127-137.
- Ciobanu, G. & Varga, C. (2010). Les compétences du terminologue raportées au Web 2.0. In *Proceedings of 6th Realiter Scientific Day*. Universidade do Algarve, Faro.
- Clancey, W. J. (1996). *Conceptual coordination: Abstraction without description*. In *International Journal of Educational Research* 27(1): pp. 5-19. Acedido em 14 de Outubro de 2009 em: <http://cogprints.org/41/1/145.htm>
- Conceição, M. C. (1999). "Terminologie, Connaissances et Industrie". in *Actes de la Conference sur la Coopération dans le Domaine de la Terminologie en Europe*. Acedido em 4 de Abril de 2011 em: www.unilat.org/dtil/aet/actes/CONCEIÇÃO.htm
- Conceição, M. C. (2005). *Concepts, Termes et Reformulations*. Presses Universitaires de Lyon.
- Condamines, A. & Aussenac-Gilles, N. (2007). *Corpus et terminologie*. R.T. Pédaque (ed.): *La redocumentarisation du monde*. Toulouse: Cepadues Editions, pp.131-147. Acedido em 14 de Setembro de 2011 em: <http://w3.erss.univlse2.fr:8080/index.jsp?perso=acondami&subURL=index.html>
- Corcho, O., Fernandez-Lopez, M., Gomez-Perez, A. (2003). Methodologies, tools and languages for building ontologies: where is their meeting point?, *Data & Knowledge Engineering*, 46(1), pp. 41 – 64.
- Cornejo, M., (2003). Utility, value and Knowledge Communities. *Knowledge Communities (to be published)*. Acedido em 4 de Maio de 2011 em: http://www.providersedge.com/docs/km_articles
- Cortese, Giuseppina, (2007). LSP: Multilingual deficiency, multicultural ambiguity. In *Evidenced-based LSP – Translation, Text and Terminology*. Ahmad Khurshid & Margaret Rogers (Eds). Peter Lang, Bern.

- Costa, R. (2001). *Pressupostos teóricos e metodológicos para a extracção automática de unidades terminológicas multilexémicas*. Tese de Doutoramento. Faculdade de Ciências Sociais e Humanas da Universidade Nova de Lisboa, Lisboa.
- Costa, R. (2004). *Corpus de spécialité: une question de types ou de genres de textes ou de discours, De la mesure des mots*. Actas do Colóquio em Homenagem a Philippe Thoiron. Université Lumière 2, Lyon.
- Costa, R. (2004). Denomination in texts for special purposes. Rahuvaeheline terminoloogiakonverents “Esti oskuskeel 2003”. Esti Keele Sihtasustus, Tallinn, Estónia.
- Costa, R. *et al.* (2004). Proceedings of the IV International Conference on Language Resources and Evaluation, LREC 2004. ELRA, Lisbon.
- Costa, R. *et al.* (2004). Proceedings of the Workshop on Computational and Computer-Assisted Terminology. ELRA, Lisbon.
- Costa, R. *et al.* (2005). Entre langue générale et langue de spécialité. Une question de collocations. Étude de Linguistique Appliquée – ELA. Paris: Didier, Érudition.
- Costa, R. (2005). Terminology, *corpus* linguistic and ontology. Proceedings of the 7th Terminology and Knowledge Engineering Conference – TKE 2005. Copenhaga: GTW.
- Costa, R. (2006). Plurality of Theoretical Approaches to Terminology. Modern Approaches to Terminological Theories and Applications. Heribert Picht [ed.] Serie: Linguistic Insights. Studies in Language and Communication. Vol.36. Berlin - Bern: Peter Lang Verlag
- Costa, R. (2006). Terminology, *Corpus* Linguistics and Ontology. In *Contrastive Studies and Valency*. Studies in Honor of Hans Ulrich Boas. Petra C. Steiner, Hans C. Boas. Stefan Scheirholz [eds.]. Peter Lang Verlag, Berlin – Bern.
- Costa, R. (2009). *O Ensino da Ciência e da Tecnologia em Português: uma questão de Terminologia*. Viseu.
- Costa, R. (2010). *Les corpus de spécialité: entre le terminologue et l’expert*. Forthcoming.
- Dam, Helle V., Engberg, Jan, Gerzymisch-Arbogast, Heidrun, (2005). *Knowledge Systems and Translation*. Mouton de Gruyter. Berlin.

- Davenport, T. H. & Prusak, L. (2000). *Working knowledge: How organizations manage what they know*. Harvard Business School Press, Boston.
- Davis, K. (2001). *Deconstruction and Translation*. St. Jerome Publishing, Manchester, UK & Northampton.
- Declerck T., A. Gomez-Perez, O. Vela, Z. Gantner, and D. Manzano-Macho. Multilingual lexical semantic resources for ontology translation. Em Proceedings of LREC 2006.
- Declerck, T., Krieger, Hans-U., et al. (2010). Ontology-based Multilingual Access to Financial Reports for Sharing Business Knowledge across Europe. Acedido em 2 de Março de 2011 em: http://www.conference2010.ia-manager.org/stories/conference2010/attachments/article/57/Declerck_MONNET-MONTIFIC.pdf
- Depecker, Loïc. (2002). *Entre signe et concept. Eléments de Terminologie générale*. Presses Sorbonne Nouvelle, Paris. pp 198. ISBN 2878542347.
- Despres, S., Szulman, S. (2008). Réseau terminologique versus Ontologie. In *Terminologie & Ontologie: Theories et Applications*. Actes de la deuxième conférence TOTh – Annecy-5 et 6 juin. France.
- Diki-Kidiri, Marcel. (1999). Le signifié et le concept dans la dénomination. *Meta : journal des traducteurs / Meta: Translators' Journal*. Vol. 44, n° 4, 1999, p. 573-581. <http://id.erudit.org/iderudit/002566ar>
- Dubuc, Robert (1987). *Manuel pratique de terminologie*. Brossard, Linguatex.
- Dunne, Keiran J. (2006). *Perspectives on Localization*. Amsterdam/Philadelphia: John Benjamins.
- Ekholm, A. (2005). *ISO 12006-2 and IFC – Prerequisites for Coordination of Standards for Classification and Interoperability*. In ITcon Vol. 10 (2005), Ekholm, pp. 275. Acedido em 2 de Maio de 2010, em: <http://www.itcon.org/2005/19/>
- Espinoza M., Gomez-Perez, A., and Mena, E. (2008). Enriching an ontology with multilingual information. In Proc. of 5th European Semantic Web Conference (ESWC'08), Tenerife, Spain.

- Espinoza M., Gomez-Perez, A., Mena, Eduardo. (2008). *LabelTranslator - A Tool to Automatically Localize an Ontology*. In Proceedings of ESWC'08 Proceedings of the 5th European semantic web conference on The semantic web: research and applications. pp. 792-796. Springer-Verlag. Berlin, Heidelberg.
- Espinoza M., Gomez-Perez, A., and Montiel-Ponsoda, H. (2009). *Multilingual and Localization Support for Ontologies*. In European Semantic Web Conference (ESWC2009). Acedido em 8 de Maio de 2010, em: <http://sid.cps.unizar.es/PUBLICATIONS/POSTSCRIPTS/eswc09-localization.pdf>
- Esselink, Bert (2000). A Practical Guide to Localization. Volume IV. John Benjamins Publishing Company, Amsterdam/Philadelphia.
- Euzenat, J., Ferrara, A., Hollink, L., Isaac, A., Joslyn, C., Malaisé, V., *et al.* (2009). Results of the ontology alignment evaluation initiative 2009. Em Workshop on Ontology Matching (OM09) at the ISWC Conference.
- Faber, Pamela. (2002). ONCOTERM: sistema bilingüe de información y recursos oncológicos. Em *La traducción científico-técnica y la terminología en la sociedad de la información*, ed. Amparo Alcina and Silvia Gamero, pp. 177-188. Castelló: Universitat Jaume I.
- Faber, P. (2009). The cognitive shift in terminology and specialized translation. Em *Monographs in Translation and Interpretation 1. A (Self-)Critical Perspective of Translation Theories*. Vidal Claramonte, Carmen África & Javier Franco Aixelá (Eds.). MonTI 1 (2009). ISSN: 1889-4178.
- Faber, P. (2010). The dynamics of specialized knowledge representation: simulational reconstruction or the perception-action interface. In *Terminology 17*, no. 1 (January 1), pp 9-29.
- Falk, I., Cruz-Lara, S., Bellalem, N., Bellalem, L. (2009). Linguistic Information for Multilinguality in Sembysem Project. Em IC3K - International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management.
- Felber, H. (1987). *Manuel de Terminologie*. Unesco e Inforterm, Paris.

- Felber, H. (2001). *Allgemeine Terminologielehre, Wissenlehre und Wissentechnik: Theoretische Grundlagen und philosophische Betrachtungen*. TermNet Publisher, Wien. pp. 283. ISBN 03-901010-13-0.
- Fernandez, S. (2007). *La poliedricidad de concepto especializado a través de la variación denominativa: primera aproximación*. Barcelona: Universidad Pompeu Fabra, Instituto Universitario de Linguística Aplicada (IULA).
- Francopoulo, G, George, M., Calzolari, N., M. Monachini, N. Bel, M. Pet, C. Soria. (2006). *Lexical Markup Framework (LMF)*. Proceedings of LREC 2006.
- Francopoulo, G., Bel, N., Georg, M., Calzolari, N., Monachini, M., Pet, M., and Soria, N. (2007). *Lexical Markup Framework: ISO standard for semantic information in NLP lexicons*. In Proceedings of the Workshop of the GLDV Working Group on Lexicography at the Biennial Spring Conference of the GLDV.
- Friedman, J. (1995). Global system, globalization and the parameters of modernity, in: Mike Featherstone, Scott Lash & Roland Robertson (eds.), pp 69-90.
- Fry, D. (2003). *The Localization Industry Primer*. Féchy: SMP Marketing and the Localization Industry Standard Association.
- Fu, B., Brennan, R., O'Sullivan, D. (2009). Multilingual Ontology Mapping: Challenges and a Proposed Framework. Workshop on Matching and Meaning 2009. Artificial Intelligence and Simulation of Behaviour Convention. Edinburgh, UK.
- Fu, B., Brennan, R., O'Sullivan, D. (2010). *Cross-Lingual Ontology Mapping and its use on the Multilingual Semantic Web*. In 1st Workshop on the Multilingual Semantic Web, at the 19th International World Wide Web Conference (WWW 2010). Raleigh, USA, April 27th, 2010, Buitelaar P., Cimiano P., Montiel-Ponsoda E., CEUR Vol.571, 2010, pp 13 – 20.
- Galinski, C., e Picht, H. (2001). Graphic and other semiotic forms of knowledge representation in Terminology management. In Wrisht, Sue E.; Budin, Gerhard, comps. *Handbook of Terminology Management*, Vol. 1, pp 42-61. John Benjamins Publishing Company, Amsterdam/Philadelphia. ISBN 1556195087.

- Gamper, J., Nejd, W., e Wolpers, M. (1999). Combining Ontologies and Terminologies in Information Systems. In Proceedings TKE 99.
- Genesereth, M. R. & Nilsson, N. J. (1987). *Logical Foundations of Artificial Intelligence*. Morgan Kaufmann, Los Altos, CA.
- Gödert, W. (2008). Ontological spine, localization and multilingual Access: Some reflections and a proposal. *New Perspectives on Subject Indexing and Classification in an International Context*. International Symposium in Honour of Magda Heiner-Freiling.
- Goguen, J. (2005). What is a Concept? F. Dau, M-L. Mugnier, G. Stumme (Eds.): ICCS 2005, LNAI 3596, Springer-Verlag Berlin Heidelberg, pp. 52-77.
- Gómez-Gauchía, H., B. Díaz-Agudo, *et al.* (2004) Two-layered approach to knowledge representation using conceptual maps and description logics. First Int. Conference on Concept Mapping, Pamplona, Spain. Acedido em 30 de Novembro de 2011, em: <http://cmc.ihmc.us/papers/cmc2004-205.pdf>
- Gouadec, D. (2003). Mondialisation, internationalisation et Normes comptables. In Colloque International de l'IALE/IALB Conference - Actes du colloque International à l'Université de Rennes 2 (septembre 2003). Paris, La Maison du Dictionnaire. Acedido em 30 de Novembro de 2011, em: http://www.colloque.net/archives/2003/ucr203_vol2.htm
- Gracia, J., Montiel-Ponsodo, E., Cimiano, P., Gómez-Pérez, A., Buitelaar, P., McCrae, J. (2012). Challenges for the Multilingual Web of Data. In *Web Semantics: Science, Services and Agents on the World Wide Web*. Volume 11, March, 2012. pp. 63-71.
- Greimas, A. J., Courtés, J. (1979). *Sémiotique. Dictionnaire raisonné de la théorie du langage*. Tome 1, Paris, Hachette, pp. 443.
- Grenon, P. (2008). A Primer on Knowledge Management and Ontological Engineering. In *Applied Ontology: An Introduction*. (Eds) Katherine Munn, Barry Smith. *Metaphysical Research*. Reicher, M., Seibt, J., Smith, B., Von Wachter, D. (eds.). Volume 9.
- Gresser, J. Y. (2008). Terminology and Information Science(s). Paper presented at the Information and Communication Technologies: From Theory to Applications, 2008. ICTTA 2008, Damascus.

- Gruber, T. (1993a). *A translation approach to portable ontology specifications*. Knowledge acquisition, 5(2), 199-220. Acedido em 2 de Abril de 2009, em: <http://ksl-web.stanford.edu/knowledgesharing/papers/README.html#ontolingua-intro>
- Gruber, T.R. (1993b). *Toward principles for the design of ontologies used for knowledge sharing*. Acedido em 4 de Abril de 2009, em: http://ksl-web.stanford.edu/KSL_Abstracts/KSL-93-04.html.
- Gruber, T.R. (1995). *Toward Principles for the Design of Ontologies Used for Knowledge Sharing*. International Journal of Human Computer Studies, 43(5-6), pp. 907-928.
- Guarino, N. (1997). Understanding, building, and using ontologies. International Journal of Human-Computer Studies, 46, pp. 293-310.
- Guarino, N. (1998). Formal ontology and information systems. In *Formal Ontology in Information Systems*. Edited by N. Gurino. Amsterdam, IOS Press. pp. 3-15.
- Guarino, N., Oberle, D. and Staab, S. (2009). What is an ontology? In *Handbook on Ontologies*. International Handbooks on Information Systems, 2009, Part 1, 1-17, DOI: 10.1007/978-3-540-92673-3_0.
- Guidère, M. (2004). De l'adaptation à la localisation publicitaire. Em *La Localisation problématique de formation*. Linguattech Editeur. Brossard (Québec).
- Hamon, T., e Nazarenko, A. (2002). *Structuration de terminologie: objectifs, pratiques, méthodes et évaluation*. Revue TAL. Volume 43:1. Juin.
- Hessen, J. (1987). *Teoria do conhecimento*. (trad. de António Correia) 8ª ed. Coimbra: Arménio Amado, 1987. pp. 206. (Coleção STVDIVM). ISBN 0-12-021198-X.
- Hirst, G. (2004). Ontology and the Lexicon. In S. Staab, and R. Studer (eds.) *Handbook on Ontologies and Information Systems*. Springer, Berlin.
- H-Know (2010). *Advanced Infrastructure for Knowledge Based Services for Building Restoring*. Acedido em 12 de Agosto de 2011, em: <http://www.h-know.eu/>
- Hoey, M. (1995). *Patterns of Lexis in Text*. Oxford University Press, Oxford.
- Hoft, N. (1995). *International Technical Communication – How to export Information about High Technology*. John Wiley & Sons, New York.

- Hohnhold, I. (1990). *Übersetzungsorientierte Terminologearbeit. Eine Grundlegung für Praktiker*. Stuttgart: InTra e.G.
- Holden, N. (2002). *Rethinking concepts for intercultural communication in the global Knowledge economy*. SIETAR 2002. Viena.
- Horrigan, J. B. (2001). Online communities: Networks that nurture long-distance relationships and local ties. Pew Internet and American Life Project. Acedido em 2 de Março de 2008, em: <http://www.pewinternet.org/reports/toc.asp?Report=47>
- Hovy, Eduard *et al.* (1999). Multilingual Information Management: Current Levels and Future Abilities. US National Science Foundation. Acedido em 3 de Fevereiro de 2012, em: <http://www.cs.cmu.edu/~ref/mlim/chapter1.html>
- Huang, C. (2000). Lexicon-driven Ontology and Conceptual Structure. Acedido em 4 de Outubro de 2009, em: www.ling.sinica.edu.tw/files/SelectedResearchProject5_EN.pdf
- Hubrich, J., Mengel, T., Müller, K., Jacobs, J. (2008). Improving Subject Access in Global Information Spaces — Reflections upon Internationalization and Localization of Knowledge Organization Systems (KOS). In: *New Perspectives on Subject Indexing and Classification. Essays in Honour of Magda Heiner-Freiling*. Leipzig u.a.: Deutsche Nationalbibliothek. S. 261-267
- Hutchins, W., Harold L., (1992). An introduction to machine translation. London: Academic Press, 1992. [ISBN: 0-12-362830-X]
- ISO (1994). Classification of information in the construction industry. ISO Technical Report 14177:1994(E). Geneva: International Organisation for Standardisation.
- ISO (1995). Building elements using explicit shape representation. Part 225 of ISO 10303, Project draft 9 June 1995, (ed. W Haas) Haas+Partner, Stuttgart, Germany.
- ISO/FDIS 1087-1:2000 (E/F) (2000). *Terminology work – Vocabulary – Part 1: Theory and application*. International Standards Organization.
- ISO/FDIS 1087-2:2000 (E/F) (2000). *Terminology work – Vocabulary – Part 2: Computer applications*. International Standards Organization.
- ISO/FDIS 704:2000 (E) (2000). *Terminology work – Principles and methods*. International Standards Organization.

- ISO/FDIS 704:2009 (E) (2009). *Terminology work – Principles and methods*. International Standards Organization.
- ISO (2001). *ISO Building construction – Organization of information about construction works - Framework for classification of information*. International Standards Organization.
- ISO (2003). Em ISO 12620. *Data Categories*. Terminology and other language resources. International Standards Organization. Acedido em: 9 de Setembro de 2009, em: <http://www.iso.org/iso/en/CatalogueDetailPage.CatalogueDetail?CSNUMBER=32347>
- ISO/TR 22134 (2007). *Practical Guidelines for Socioterminology*. International Standards Organization.
- Jackendoff, R. (1995). *Languages of the Mind – Essays on Mental Representation*. MIT Press. London.
- Jenkins, H. (2006). *Confronting the Challenges of Participatory Culture: Media Education for the 21st Century*.
- Kageura, K. (1997). “Multifaceted/Multidimensional Concept Systems”. *Handbook of Terminology Management – Volume 1: Basic Aspects of Terminology Management*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Kageura, K. (2002). *The Dynamics of Terminology: a descriptive theory of term formation and terminological growth*. John Benjamins Publishing Company. Amsterdam/Philadelphia.
- Kargioti, E., Kourtesis, D., Bibikas, Dimitris, Paraskakis, Iraklis and Boes, U. (2010). *MORMED: towards a multilingual social networking platform facilitating medicine 2.0*. Em: Bamidis, P. D. and Pallikarakis, N., (eds.) XII Mediterranean Conference on Medical and Biological Engineering and Computing 2010. XII Mediterranean Conference on Medical and Biological Engineering and Computing 2010, 27-30 May 2010, Chalkidiki, Greece. IFMBE Proceedings, 29 . Springer , Berlin / Heidelberg, pp. 971-974.
- Karsch, B. I. (2009). Profile of a Terminologist in Localization environments. Em *The Journal of Internationalization and Localization*. Volume 1. Endrik. J. Kockaert (Ed.). Lessius University College/ K. U. Leuven.

- Keiran, J. D. (2006). *Perspectives on Localization*. In American Translators Association Scholarly Monograph Series – Volume XIII. John Benjamins Publishing Company. Amsterdam/Philadelphia.
- Kelly, M. (2009). *A third space for europe: intercultural communication in european language policy*. *European Journal of Language Policy* 1(1), pp. 1-20.
- Kerényi, K. (1976). *Dionysos: Archetypal Image of Indestructible Life*. University Press. Princeton.
- Kerremans, K., Temmerman, R., and Tummers, J. (2003). *Representing Multilingual and Culture Specific Knowledge in a VAT Regulatory Ontology: Support from the Termontography Method*. [OTM Workshops 2003](#): pp. 662-674.
- Kerremans K. e T. R. (2004). Towards multilingual, termontological support in ontology engineering. Em: *Proceedings Workshop on Terminology, Ontology and Knowledge representation*. Lyon, France.
- Kerremans K., R. Temmerman & J. Tummers. (2004a). Discussion on the Requirements for a Workbench supporting Termontography. Em WilliAms, G. & S. Vessier (eds.). *Proceedings of the XIth Euralex International Congress 2004, 6-10 July, Lorient, France*. Lorient: Université de Bretagne-Sud. pp. 559-570.
- Kerremans, K., R. Temmerman e G. Zhao. (2005). Terminology and Knowledge Engineering in Fraud Detection. Em *Proceedings of the International conference on Terminology and Knowledge Engineering*. Copenhagen, 16th – 19th August 2005.
- Kharatmal, Meena & Nagarjuna, G. (2010). *Introducing Rigor in Concept Maps*. In M. Croitoru, S. Ferre, and D. Lukose (Eds.), *Lecture Notes in Artificial Intelligence: Vol. 6208. International Conference on Conceptual Structures 2010: From Information to Intelligence* (p. 199-202). Springer-Verlag. Berlin, Germany. Doi: 10.1007/978-3-642-14197-3_22
- Kostina, I. (2010). *Variación conceptual de términos en el discurso especializado*. Barcelona: Universidad Pompeu Fabra, Instituto Universitario de Lingüística Aplicada (IULA).
- L'Homme, M.-C. (2004). *La terminologie: principes et techniques*. Montréal: Les Presses de l'Université de Montréal.

- [Langacker, R. W.](#) (1987). *Foundations of cognitive grammar: Theoretical Prerequisites*. Stanford University Press. Stanford, CA.
- Laurén, C., Myking, J., Picht, H. (1998). *Terminologie unter der Lupe: vom Grenzgebiet zum Wissenschaftszweig*. TermNet. Vienna. pp. 353. ISBN 3-901010-15-7. (IITF Series 9).
- Lefevre, A. (1992). *Translation/History/Culture – A Sourcebook*. London & New York: Routledge.
- Lenci, A. (2008). The Life-Cycle of Knowledge. Em Huang, C.-R., Calzolari, N., Gangemi, A., Lenci, A., Oltramari, A., Prevot L. (eds.). *Ontologies and the Lexicon A Natural Language Processing Perspective*. Cambridge University Press. Cambridge.
- Lerat, P. (2009). *Communautés de travail et diversité terminologique*. Communication au colloque LTT (Lisbonne, octobre 2009). À paraître dans les actes.
- Lino, M. T. (1995). Da Constituição de *Corpora* à Lexicografia Informatizada de Especialidade. Em *Actas do XI Encontro Nacional da Associação Portuguesa de Linguística* (vol. II). Lisboa. APL, pp. 67-92.
- LMF-Lexical Markup Framework (2006). Em ISO 24613. Language Resource Management. Acedido a 13 de Março de 2011, em: http://lirics.loria.fr/doc_pub/LMF%20rev9%2015March2006.pdf
- López, M. F., Gómez-Pérez, A., Sierra, J. P. (1999). *Building a Chemical Ontology Using Methontology and the Ontology Design Environment*. Intelligent Systems and their Applications, IEEE.
- Lustig, M. W. e Koester, J. (2006). *Intercultural competence: interpersonal communication across cultures*. Pearson.
- Martins, J. P. (2006). *Modelação do Fluxo de Informação no Processo de Construção*. Tese Doutorado. Faculdade Engenharia Universidade do Porto. Acedido a 17 de Março de 2011, em: <http://paginas.fe.up.pt/~jppm/docs/Tese.pdf>
- Meyer, I. (1992). Knowledge Management for Terminology-Intensive Applications: Needs and Tools. Em *Lexical Semantics and Knowledge Representation*. Eds. James Pustejovsky and Sabine Bergler. Berlin, Springer Verlag, pp. 21-37.

- Meyer, Ingrid *et al.* (1997). Systematic Concept Analysis within a Knowledge-Based Approach to Terminology. Em *Handbook of Terminology Management – Volume 1: Basic Aspects of Terminology Management*. John Benjamins Publishing Company. Amsterdam/Philadelphia.
- Meyer, I. (2001). Extracting knowledge-rich contexts for terminography. In *Recent Advances in Computational Terminology*. Bourigault, D., C. Jacquemin and M.-C. L'Homme (eds.). Amsterdam/Philadelphia: John Benjamins, pp. 279-302.
- Michalski, R. (1993). Beyond Prototypes and Frames: the two-tiered Concept Representation. Em Van Mechelen, I. Et alii, (ed.), *Categories and Concepts, Theoretical Views and Inductive Data Analysis*. London: Academic Press.
- Mondary, T., Després, S., Nazarenko, A., and Szulman. S. (2008). *Construction d'ontologies à partir de textes: la phase de conceptualisation*. 19es Journées Francophones d'Ingénierie des Connaissances (IC 2008), Nancy : France
- Monteiro, P. M. (1998). *Classificação da informação na indústria da construção*. Dissertação de Mestrado. Faculdade Engenharia. Universidade do Porto.
- Montiel-Ponsoda, E., Peters, W., Aguado de Cea, G., Espinoza, M., Pérez, A. G., e Sini, M. (2008). *Multilingual and localization support for ontologies*. Technical report, D2.4.2 NeOn Project Deliverable.
- Montiel-Ponsoda, E., Aguado de Cea, G., Espinoza, M., Pérez, A. G., Peters, W. (2008a). *Modelling Multilinguality in Ontologies*. [COLING \(Posters\) 2008](#): 67-70.
- Montiel-Ponsoda, E. (2011). *Multilingualism in Ontologies Multilingual Lexico-Syntactic Patterns for Ontology Modeling and Linguistic Information Repository for Ontology Localization*. Tese de Doutoramento. Universidad Politécnica de Madrid. E.T.S.I. Montes
- Nakata, K. (2007). *A Grounded and Participatory Approach to Collaborative Information Exploration and Management*. In Proceedings of HICSS '01 Proceedings of the 34th Annual Hawaii International Conference on System Sciences. (HICSS-34)-Volume 1 - Volume 1.
- Newmark, Peter. (1988). *A Textbook of Translation*. Prentice Hall, Cambridge.

- Newmark, Peter. (2007). *A new theory of Translation*. Consultado a 23 de Junho de 2012, em: [http://www.phil.muni.cz/plonedata/wkaa/BSE/BSE_2007-33_Offprints/BSE%202007-33%20\(101-114\)%20Newmark.pdf](http://www.phil.muni.cz/plonedata/wkaa/BSE/BSE_2007-33_Offprints/BSE%202007-33%20(101-114)%20Newmark.pdf).
- Nida, E. (2000). Principles of Correspondence. Em *The Translation Studies Reader*. Ed. Lawrence Venuti. London: Routledge, 126 – 141.
- Nirenburg, Sergei, & Raskin, V. (2004). *Ontological Semantics*. Cambridge, MA: MIT Press.
- Novak, Joseph D. (2003). *The theory underlying concept maps and how to construct them*. Consultado a 23 de Maio de 2012, em: <http://cmap.coginst.uwf.edu/info>
- O'Hagan, M., Ashworth, D. (2002). *Translations-mediated communication in a digital world. Facing the challenges of Globalization and Localization*. Clevedon. Multilingual Matters Ltd.
- Otman, G. (1996). *Les Représentations Sémantiques en Terminologie*. Paris: Sciences Cognitives, Mason.
- Pagans, M. (2002). A localização: Localização, localizamo-nos? In *Revista Tradumática*, nº 1. Centro virtual Camões. Consultado a 12 de Fevereiro de 2012, em: <http://cvc.instituto-camoes.pt/tradumatica/rev1/mpagansPT.html>
- Pazienza, M. T. e Stellato, A. (2006). *Exploiting linguistic resources for building linguistically motivated ontologies in the semantic web*. Em Second Workshop on Interfacing Ontologies and Lexical Resources for Semantic Web Technologies (OntoLex2006), held jointly with LREC2006, May 24-26, 2006, Genoa, Italy.
- Pazienza, M. T. e Stellato, A. (2006a). [*An Environment for Semi-automatic Annotation of Ontological Knowledge with Linguistic Content*](#). Em 3rd European Semantic Web Conference (ESWC 2006) Budva, Montenegro, June 11-14, 2006.
- Pazienza, M. T. e Stellato, A. (2006b). [*An open and scalable framework for enriching ontologies with natural language content*](#). In The 19th International Conference on Industrial, Engineering & Other Applications of Applied Intelligent Systems (IEA/AIE'06), special session on Ontology & Text Annecy, France, June 27-30, 2006.
- Pearson, J. (1998). *Terms in Context*. John Benjamins Publishing Company . Amsterdam/Philadelphia.

- Pedersen, T., Banerjee, S., Patwardhan, S. (2005). *Maximizing Semantic Relatedness to Perform Word Sense Disambiguation*. Research Report UMSI 2005/25, University of Minnesota Supercomputing Institute.
- Pereira, C., Silva, M., Fernandes, J., Soares, A. L. (2007). *Understanding users' response to ontology based systems in the context of an Enterprise Sponsored Virtual Community*. 8th IFIP Working Conference on VIRTUAL ENTERPRISES - PRO-VE 2007. Universidade do Minho. Guimarães.
- Pereira, C., Soares, A. L. (2008). *Ontology development in collaborative networks as a process of social construction of meaning*. Em *On the Move to Meaningful Internet Systems: OTM 2008 Workshops, Lecture Notes in Computer Science, Springer Berlin / Heidelberg, November, 2008*.
- Pereira, C., Sousa, C., Soares, A. L. (2009). *A socio-semantic approach to collaborative domain conceptualization*. Em *On the Move to Meaningful Internet Systems: OTM 2009 Workshops, Lecture Notes in Computer Science, Springer Berlin /Heidelberg*.
- Pereira, C. (2010). *Organização da informação e conhecimento em redes colaborativas como um processo de construção social do significado: uma teoria e um método prático*. Tese de Doutoramento. Faculdade de Engenharia da Universidade do Porto.
- Pereira, C., Sousa, C., e Soares, A. L. (2012). *A socio-semantic approach to support conceptualisation processes: a case study in an R&D project*. Accepted for publication in the *International Journal of Computer Integrated Manufacturing*.
- Peters, W., Vossen, P., Díez-Orzas, P., Adriaens, G. (1998). *Cross-linguistic Alignment of Wordnets with an Inter-Lingual-Index*. pp. 149–179.
- Peters, W., Montiel-Ponsoda, E., e Aguado de Cea, G. (2007). *Localizing ontologies in OWL*. In *OntoLex 2007*.
- Picht, H. & Draskau, J. K. (1985). *Terminology: An Introduction*. Surrey: University of Surrey.
- Pinto, H. S., e Martins, J. P. (2001). *Ontology Integration: How to perform the Process*. *Proceedings of the International Conference on Knowledge Capture*. Victoria, British Columbia, Canada. Consultado a 5 de Junho de 2012, em: <http://www-agki.tzi.de/buster/IJCAIwp/Finals/pinto.pdf>

- Prandi, M. (2011). *Signes, signifiés, concepts : pour un tournant philosophique en linguistique*. TOTH 2011.
- Prévot, L., [Chang](#), C., Desalle, Y. (2010). Computational Modeling of Verb Acquisition, from a Monolingual to a Bilingual Study. [PACLIC](#): 841-851.
- Pym, A. (2002). *Redefining Translation Competence in an Electronic Age. In defense of a minimalist approach*. Tinet. Consultado a 17 de Junho de 2012, em: <http://www.tinet.cat/~apym/on-line/training/competence.pdf>
- Pym, A. (2004). *The Moving Text, Localization, Translation, and Distribution*. John Benjamins Publishing Company. Amsterdam/Philadelphia.
- Pym, A. (2004a). *Localization from the Perspective of Translation Studies: Overlaps in the Digital Divide?* Paper presented to the [SCALLA conference](#), Kathmandu, January 2004. Consultado a 14 de Maio de 2011, em: <http://www.elda.org/en/proj/scalla/SCALLA2004/Pymv2.pdf>
- Pym, A. (2006). *Localization, Training, and the Threat of Fragmentation*. Tinet. Consultado a 30 de Abril de 2011, em: <http://www.tinet.cat/~apym/on-line/translation/translation.html>
- Quirion, J. (2003). *La mesure de l'implantation terminologique*. Étude terminométrique du domaine des transports au Québec, Langues et sociétés, N0.40.
- Rastier, F. (2001). *Arts et sciences du texte*. Paris. PUF.
- Rastier, F. (2004). *Enjeux épistémologiques de la linguistique de corpus*. Consultado a 20 de Janeiro de 2012, em: <http://www.revue-texto.net/index.php?id=543>
- Remígio, A. R. (2010). *Processo terminográfico: vertentes conceptual, comunicativa e textual*. Tese de Doutoramento. Universidade de Aveiro.
- Rey, A. (1979). *La Terminologie: noms et notions*. Paris: Presses Universitaires de France. 127 p. ISBN 2130445284.
- Rey, A. (1995). *Essays on Terminology*. John Benjamins Publishing Company. Amsterdam/Philadelphia.
- Roche, C. (2003). [Ontology: a Survey](#). In 8th Symposium on Automated Systems Based on Human Skill and Knowledge, IFAC, September 22-24. Göteborg, Sweden.

- Roche, C. (2005). Terminologie et ontologie. In *Langages*. 39e année, n°157, 2005. pp. 48-62. Persee.
- Roche, C. (2006). *Ontology for Long-Term Knowledge Roche*. In [Lecture Notes in Computer Science](#). Volume 4031/2006, 583-589. Springer. (PAIS)
- Roche, C. (2006a). *How words map concepts*. VORTE 2006. EDOC Conference, Hong Kong.
- Roche, C. (2006b). *Lexical and Conceptual Structures in Ontology*. . In *Advances in Applied Artificial Intelligence - Lecture Notes in Computer Science* (pp. 1034-1041): Springer Berlin / Heidelberg.
- Roche, C. (2007). *Le terme et le concept: fondements d'une ontoterminologie*. Paper presented at the Terminologie & Ontologie: Théories et Applications, Annecy.
- Roche, C. (2007b). *Dire n'est pas concevoir*. IC 2007: 18e Journées Francophones d'Ingénierie des Connaissances. Grenoble.
- Roche, C. (2008). *Faut-il revisiter les Principes terminologiques?*. Conférence TOTH 2008 Terminologie & Ontologie: Théories et Applications. Annecy 5 et 6 juin 2008 – pp. 53-72.
- Roche, C., and Costa, R. (2011). Prefácio à conferência TOTH 2011. Toth 2011.
- Rogers, M. (2007). *Terminological equivalence: Probability and consistency in technical translation*. EU-High-Level Scientific Conference Series MuTra 2007 – LSP Translation Scenarios: Conference Proceedings.
- Rogers, M. (2007a): Lexical chains in technical translation: A case study in indeterminacy. Em B. Antia (ed.): *Indeterminacy in LSP and Terminology: Studies in Honour of Heribert Picht*. Amsterdam/Philadelphia: John Benjamins Publishing Company. pp. 15-35.
- Rogers, M. (2008). Terminological Equivalence: Probability and Consistency. Em *LSP Translation Scenarios. Selected Contributions of the EU Marie Curie Conference Vienna 2007*. Mutra Journal · Volume 02. Heidrun Gerzymisch-Arbogast & ATRC Group.
- Romualdo, C. (2010). *Funções e competências do terminólogo nas Empresas Luso-polacas*. Realiter.
- Rondeau, G. (1984). *Introduction à la Terminologie*. Québec: Gaetan Marin Editeur.

- Rothkegel, A. (2003). *Knowledge and text types*. Em Knowledge systems and translation. Helle V. Dam, Jan Engberg. Heidrun Gerzymisch-Arbogast.
- Sager, J. C. (1990). *A Practical Course in Terminology Processing*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Sager, J. (1994). *Language Engineering and Translation – Consequences of automation*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Sánchez, D. M., Cavero, J. M. e Marcos, E. (2005). *On Models and Ontologies*. In Proceedings of the 1st International Workshop on Philosophical Foundations of Information Systems Engineering (PHISE'05), Porto, Portugal.
- Sánchez, D. M., Cavero, J. M. e Martínez, E. M. (2007). The Road Toward Ontologies, Em *Ontologies: A Handbook of Principles, Concepts and Applications*. In Information Systems, Sharman R., Kishore R. and Ramesh R., Springer US, 3-20.
- Sandrini, P. (2005). *Website Localization and Translation*. In *MuTra 2005 – Challenges of Multidimensional Translation: Conference Proceedings*.
- Santos, C. (2010). *Terminologia e Ontologias: Metodologias para Representação do Conhecimento*. Tese de Doutorado. Universidade de Aveiro.
- Sapir, E. (1929/1964). *The status of linguistics as a science*. *Language*, 5, pp. 207–214. Reprinted in: Mandelbaum, David G. (ed.), *Culture, Language, and Personality: Selected essays of Edward Sapir*, University of California Press, Berkeley.
- Savourel, Y. (2001). *XML: Internationalization and Localization*. Indianapolis: Sams.
- Schäler, R., (2007). Localization. Em: *Encyclopedia of Translation Studies*. Baker, M. and Saldanha, G. (Eds.), second edition, pp. 157-161.
- Schalley, A. C. e Zaefferer, D. (2007). Ontolinguistics - an outline. Em *Ontolinguistics. How Ontological Status Shapes the Linguistic Coding of Concepts*, ed. Andrea C. Schalley and Dietmar Zaefferer. (Trends in Linguistics. Studies and Monographs 176.) Berlin/New York: Mouton de Gruyter.
- Schmitz, K. (2007). Indeterminacy of terms and icons in software localization. Em: Antia, Bassey (Ed.): *Indeterminacy in LSP and Terminology. Studies in Honour of Heribert Picht*. Amsterdam/Philadelphia: John Benjamins Publishing Company, S. pp. 49-58.

- Schütze, Hinrich (1997). *Ambiguity Resolution in Language Learning – Computational and cognitive models*. Stanford: CSLI Publications.
- Silva, M., Simões, D., Costa, R. e Lucas, A. (2005). *Know-Construct: Knowledge based community management in the Construction Industry*. In Proceedings of the 7th Terminology and Knowledge Engineering Conference, Workshop Working with Terminology and Knowledge Management Systems. Copenhaga: GTW.
- Silva, M. (2007). Termo e imagem: A representação na conceptualização e transmissão do conhecimento. *3Ts – Revista de Tradução Terminologia e Tecnologia*. Universidade de Aveiro.
- Silva, M., Costa, R., Soares, A. L. (2011). Ontologias de Domínio. Em *Colloque Terminologie, Rédaction Technique: des ponts entre le français et le portugais*. L'Université. Paris 3 Sorbonne. França.
- Silva, M., Costa, R., Soares, A. L. (2012). Multilingual Ontology Specification: a collaborative approach. Em *Proceedings of Terminology and Knowledge Engineering Conference 2012 - TKE 2012*. Madrid. Spain.
- Silva, M., Costa, R., Soares, A. L. (2012). Supporting collaboration in multilingual ontology specification: the conceptME approach. Em Workshop Challenges to knowledge representation in multilingual contexts. (Eds. Rute Costa, Manuel Silva, António Lucas Soares). TKE 2012
- Slodzian, M. (2000). L'émergence d'une terminologie textuelle et le retour du sens. In *Le Sens en Terminologie*. Lyon: Presses Universitaires de Lyon.
- Smith, B. (1998). The basic tools of formal ontology. In N. Guarino (Ed.), *Formal Ontology in Information Systems* (Vol. Frontiers in Artificial Intelligence and Applications, pp. 19-28). Amsterdam: IOS Press.
- Snell-Hornby, M. (2000). *Communicating in the Global Village: On Language, Translation and Cultural Identity*, em Christina Schäffner (ed.) *Translation in the Global Village* (pp. 11-28) Clevedon: Multilingual Matters.

- Soares, A. L., Sousa, J., (2004). *Modeling Social Aspects of Collaborative Networks*. Em Camarinha *et al.* (Eds.) *Collaborative Networked Organizations*. Amsterdam, Kluwer Academic Press.
- Soares, L., Simões, D., Silva, M., Madureira, R. (2006). *Developing Enterprise Sponsored Virtual Communities: The case of a SME's Knowledge Community*. Em *On the Move to Meaningful Internet Systems 2006: OTM 2006 Workshops*. (Part 1). Springer Verlag. Berlin. pp. 269-278.
- Soeiro, A. (2010). *E-Training Engineers in Rehabilitation of Buildings and Knowledge Management: Project H-Know*. 12th IACEE World Conference on Continuing Engineering Education. Singapore.
- Sommer, H. (2003). *Computers and Translation. A translator's guide*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Sousa, C., Pereira, C., Soares, A. L. e Costa, R. (2012). Supporting the identification of conceptual relations in semi-formal ontology development. (Eds. Soares, A. L., Costa, R., Pereira, C.). Em *Proceedings of the Workshop on Collaboration in Terminology and Knowledge Application*. LREC, 2012. Istanbul. ELRA. Consultado a 22 de Março de 2012, em: <http://www.lrec-conf.org/proceedings/lrec2012/workshops/08.LREC%202012%20Terminology%20Proceedings.pdf>
- Sousa, C., Pereira, C., e Soares, A. L.. (2012a). *Discussing and collaborating through concepts: the conceptME approach*. KEOD 2012. Forthcoming.
- Sousa, H. (2001). *Caracterização da actividade construção. Capítulo 2 dos apontamentos da disciplina de Gestão de projectos da licenciatura em Eng^a Civil da FEUP*. FEUP, Porto.
- Sprung, C. R. (2000). *Translating into Success*. Cutting-Edge Strategies for Going Multilingual in a Global Age American Translators Association, Scholarly Monograph Series. Vol XI. Amsterdam: John Benjamins Publishing Company.
- Staab, S., Studer, R. (2009). *Handbook on Ontologies*. International Handbooks on Information Systems. Staab, Steffen; Studer, Rudi (Eds.), 2nd ed., Springer.

- Studer, R., Benjamins, R., Fensel, D. (1998). Knowledge engineering: Principles and methods. In *Data & Knowledge Engineering*, 25(1–2): pp. 161–198.
- Sowa, J. (2000). *Knowledge Representation: Logical, Philosophical, and Computational Foundations*. Pacific Grove, CA: Brooks/Cole Publishing Co.
- Sowa, J. (2006). Building, Sharing, and Merging Ontologies. Consultado a 19 de Janeiro de 2012, em: <http://www.jfsowa.com/ontology/ontoshar.htm>.
- Suárez-Figueroa, M. C., and Gómez-Pérez, A. (2008). Towards a Glossary of Activities in the Ontology Engineering Field. Em *Proceedings of the 6th Language Resources and Evaluation Conference (LREC08)*. Marrakech (Morocco).
- Taboada, M. *et al.* (2003). Modelling knowledge bases by reusing generic ontologies and unified terminology servers, *International Journal of Knowledge-Based Intelligent Engineering Systems*, 7 (4): pp. 190-197. Consultado a 10 de Março de 2012, em: <http://aiff.usc.es/~elchus/personal/download/taboadakES.pdf>.
- Tallone, L. (2010). *Tradutores: até que a localização nos separe?* In *Polissema N 10*. ISCAP. Porto.
- Temmerman, R. (2000). *Towards new ways of terminology description: the sociocognitive approach*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Texin, T. (2003). *The Challenges of Defining locale Structures: Language and culture identifiers need to be adaptable*. In *Multilingual Computing & Technology*, Vol. 14, No 55, Issue 3, Sandpoint (Idaho), Multilingual. Computing Inc, pp. 24-28.
- Tricot, C. and Roche, C. (2006). *Visualisation of Ontology: a focus and context approach*. InSciT2006, Mérida, Spain.
- TMF - Terminological Markup Framework. (2003). In ISO 16642, Computer applications in terminology. Consultado a 7 de Março de 2012, em: <http://www.loria.fr/projets/TMF/>
- Unesco. (2005). *Guidelines for Terminology Policies: Formulating and implementing terminology policy in language communities*. Paris. Infoterm.
- Uschold, M. e Gruninger, M. (1996). *Ontologies: principles, methods and applications*. In *The Knowledge Engineering Review*, 11(2), pp. 93-136.

- Vieu, L. (2011). On the ontological coherence of lexical resources. In *Workshop Proceedings of the 9th International Conference on Terminology and Artificial Intelligence TIA2011*. Consultado a 7 de Fevereiro de 2012, em: <http://tia2011.crim.fr/Workshop-Proceedings/pdf/TIAW12.pdf>.
- von Hahn, W. e Vertan, C. (2005). Challenges for the Multilingual Semantic Web. In *Proceedings of the Workshop on Semantic Web Technologies for Machine Translation. MT Summit X Workshop*. Phuket, Thailand. Consultado a 11 de Novembro de 2011, em: <http://www.mt-archive.info/MTS-2005-SW-WS.htm>.
- Vossen, P. (1997). *Eurowordnet: a multilingual database for information retrieval*. In *Proceedings of the workshop on Cross-language Information Retrieval*, Zurich. Amsterdam – Philadelphia: John Benjamins Publishing Company.
- Vossen, P. (1998). *EuroWordNet: A Multilingual Database with Lexical Semantic Networks*. Norwell, MA, USA: Kluwer Academic Publishers.
- Whorf, B. L. (1940). *Science and Linguistics*. *Technology Review* 42(6): pp. 229-31, 247-8. Also in B. L. Whorf (1956): *Language, Thought and Reality* (ed. J. B. Carroll). Cambridge, MA: MIT Press.
- Warburton, K. (2010). *Language resources and their commercial applications*. Consultado a 7 de Outubro de 2011, em: <http://www.saltcymru.org/wordpress/?p=722&lang=en>.
- Wills, W. (1996). *Knowledge and Skills in Translator Behaviour*. Amsterdam & Philadelphia: John Benjamins Publishing Company.
- Wright, S. E., Gerhard, B. (2001). *Handbook of Terminology Management*. Amsterdam – Philadelphia: John Benjamins Publishing Company.
- Wright, S. E. (2007). Coping with indeterminacy – Terminology and knowledge representation resources in digital environments. Em: Antia, Bassey (Ed.): *Indeterminacy in LSP and Terminology. Studies in Honour of Heribert Picht*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Wüster, E. (1985). *Einführung in die allgemeine Terminologielehre und terminologische Lexicographie*. Wien: Infoterm.

ÍNDICE DE FIGURAS

Figura 1	- Relação entre actividade de conceptualização e desenvolvimento colaborativo de ontologias	39
Figura 2	- Conceitos GILT	80
Figura 3	- Arquitectura global da base de dados do EuroWordNet	102
Figura 4	- Fluxo de trabalho do modelo Termontography	103
Figura 5	- Estrutura geral da metodologia GENOMA-KB	104
Figura 6	- Ambiente de tradução do LabelTranslator	111
Figura 7	- Abordagem típica à localização de ontologias	122
Figura 8	- Abordagem proposta à localização de ontologias	124
Figura 9	- Método para a especificação multilingue de ontologias: Visão de alto nível	128
Figura 10	- Interligação entre as ontologias em diferentes línguas	147
Figura 11	- Espaço de conceptualização e localização – acesso a definições	152
Figura 12	- Espaço de conceptualização e localização – escolha de equivalentes	154
Figura 13	- Fluxo de trabalho do método para a especificação multilingue de ontologias	161
Figura 14	- Papel dos especialistas da rede	164
Figura 15	- Papel do especialista	166
Figura 16	- Conceito de alto nível do sistema H-Know	174
Figura 17	- Arquitectura do sistema H-Know	176
Figura 18	- Processo de construção da ontologia H-Know: da H-KnowG (global) às ontologias H-KnowL (locais)	177
Figura 19	- Processo construtivo I (<i>building lifecycle</i>)	184
Figura 20	- Processo construtivo II	184

Figura 21	- Processo construtivo III	185
Figura 22	- Modelo da Norma ISO 12006-2	194
Figura 23	- Modelo global da Ontologia H-Know	200
Figura 24	- Espaço de conceptualização e localização - Mapa <i>construction process</i>	203
Figura 25	- Espaço de conceptualização e localização: exemplo de uso	213
Figura 26	- Proposta de adição de JF _{fr}	222
Figura 27	- Proposta de adição de JPM _{pt}	223
Figura 28	- Proposta de modificação de AS _{pt}	224
Figura 29	- Proposta de adição de GH _{it}	225
Figura 30	- Proposta de adição/modificação de JF _{fr}	226
Figura 31	- Proposta de modificação de JPM _{pt} e de AS _{pt}	227
Figura 32	- Proposta de eliminação de JPM _{pt}	229
Figura 33	- Proposta de eliminação de AS _{pt}	230
Figura 34	- Proposta de eliminação de JPM _{pt}	232
Figura 35	- Proposta de adição de JPM _{pt}	233

ÍNDICE DE TABELAS

Tabela 1	- Critérios para a selecção de fontes	133
Tabela 2	- Exemplo da análise de recursos do domínio	190
Tabela 3	- Tipologia de textos do <i>corpus</i> do domínio da reabilitação	192
Tabela 4	- Mapas conceptuais constituintes da amostra	201
Tabela 5	- Tipo de relações conceptuais	201

APÊNDICES

Apêndice A: Critérios para a avaliação de fontes

<i>Sources evaluation criteria¹</i>		
<i>Type of source</i>	<i>Common Criteria</i>	<i>Description</i>
All	Origin	developer(s)
	Relevance	for the pre-defined areas of analysis for specific cultural and professional context(s)
	Adequacy	from the domain expert point of view from the ontologist point of view
	Completeness	explicit in-depth coverage
	Comprehensiveness	domains addressed in the area
	Ease of data acquisition	possibility of access and reuse (merge/integrate)
	Language	language(s) in which it is available multilingual features language independence
	Current status	finished, work in progress, in revision
	<i>Specific Criteria</i>	
Ontologies	Conceptual framework/model	ontology assumptions and ontological commitment and their relation to KC objectives
	Type of concept	identification of generic concepts and relationships identification of domain concepts and relationships
	Design principles	internal structure
	Knowledge acquisition	quality of knowledge sources adequacy of knowledge acquisition practices
	Supported applications	applications supporting the ontology codification language

¹ Silva, Manuel, Simões, D., Costa, R. e Lucas, A. (2005). *Know-Construct: Knowledge based community management in the Construction Industry*. In Proceedings of the 7th Terminology and Knowledge Engineering Conference, Workshop Working with Terminology and Knowledge Management Systems. Copenhagen: GTW.

	Documentation available	Type of documentation available and accessibility
	Consistency	consistency of the application of the relations
	Modularity	which concepts are represented in which modules
Terminologies	Terminology purpose and scope	operational terms – functions the terminology is intended to serve
	Standardized/non-standardized	implemented as standard other type
	Granularity	level of complexity of the available data
	Quality of the definitions	do they follow unified patterns, are simple, clear, concise, etc.
	Interconnectivity	to what extent is the terminology mappable to coding systems or terminologies
	Precision and recall	retrieval effectiveness
	Normalization	of content and semantics
	Responsiveness	frequency of update
Classifications	Classification purposes	classification purposes and their relation to KC objectives
	Conceptual framework	classification assumptions and their relation to KC objectives
	Classification scope	domain(s)
	Type of concepts	degree of abstraction/specificity
	Previous use	use in ontology projects and outcome analysis

Apêndice B: Guião para o especialista

Guião para o especialista

Objectivos da localização dos mapas conceptuais

No contexto de uma rede colaborativa multilingue, a representação conceptual do conhecimento de um domínio advém de uma estruturação obtida através de negociação e consenso entre os especialistas do domínio.

Nesse sentido, este processo destina-se a que:

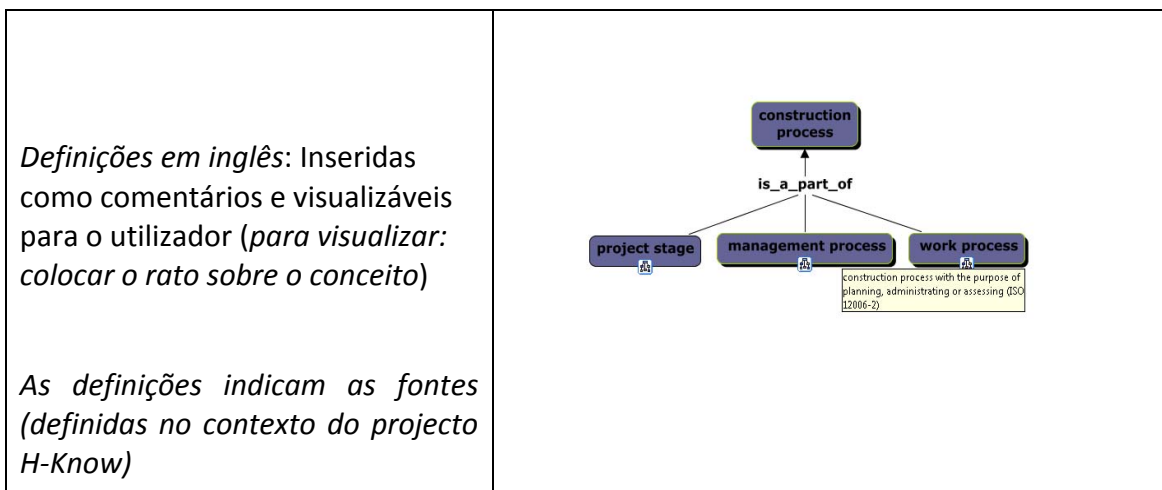
1. se desenvolvam representações conceptuais do conhecimento do domínio nas línguas dos diferentes especialistas da rede,
2. os vários especialistas participem na construção partilhada da ontologia, tendo também em conta a representação desta na sua língua e cultura de origem.

Análise e localização dos mapas conceptuais

ESPAÇO DE CONCEPTUALIZAÇÃO E LOCALIZAÇÃO:

Este espaço, baseado em mapas de conceitos é composto por diferentes elementos, que descrevemos em seguida e que pretendem auxiliar a sua participação no processo de análise dos mapas conceptuais e no processo de localização. Assim:

- Cada mapa conceptual possui um conjunto de elementos de auxílio à conceptualização e à localização do domínio, nomeadamente:

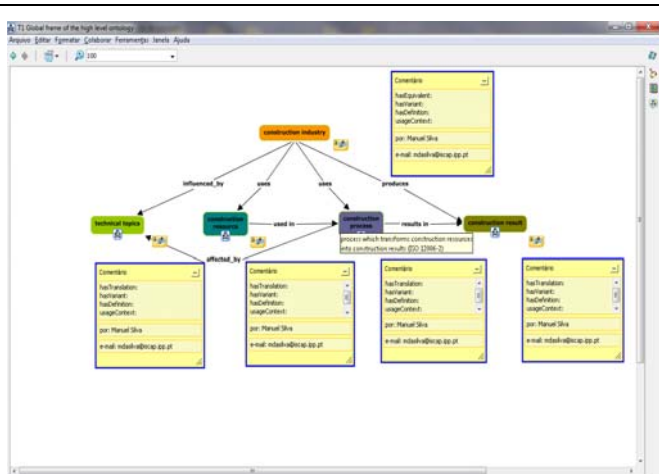


Através do uso do sistema de comentários (visível na figura seguinte) poderá inserir as suas sugestões/alterações e inserir informação como:

- a) uma ou mais variantes
- b) definições (na sua língua)
- c) contexto de uso (do termo)

Em cada caixa de comentário encontrará os seguintes elementos que poderá preencher:

1. *hasTranslation:*
2. *hasVariant:*
3. *hasDefinition:*
4. *usageContext:*



Especialista: pode inserir definições/notas na zona de comentários, junto dos equivalentes propostos para sustentar a sua escolha e, eventualmente, promover a discussão.

ELEMENTOS DE APOIO AO DESENVOLVIMENTO DA LOCALIZAÇÃO

FONTES DE PARTIDA

(Estas fontes estiveram na base da conceptualização inicial e forneceram as definições que encontrará inseridas nos mapas conceptuais.)

UNICLASS - Unified Classification for the Construction Industry	http://en.wikipedia.org/wiki/Uniclass
Thesaurus – European Heritage Network	http://www.european-heritage.net/sdx/herein/thesaurus/consult.xsp?class=1&nr=361
OmniClass Construction Classification System	http://www.omniclass.ca/
EPIC – European Product Information Co-Operation	Enviada em anexo
ISO 12006-2	Enviada em anexo

FERRAMENTAS DE APOIO À REPRESENTAÇÃO CONCEPTUAL

Os mapas conceptuais estão em formato Cmap. Deverá utilizar o CmapTools para a sua edição.

IHMC CmapTools	http://cmap.ihmc.us/download/
----------------	---

FERRAMENTAS DE APOIO À LOCALIZAÇÃO (TRADUÇÃO)

Aplicações web de **tradução automática**

Google translator	http://translate.google.pt/
Bing Translator	http://www.microsofttranslator.com/
Babel Fish	http://babelfish.yahoo.com/
Linguee (<i>corpus de exemplos</i>)	http://www.linguee.pt/

Bases de dados terminológicas:

IATE (BDT Comunidade Europeia)	http://iate.europa.eu
--------------------------------	---

Bases de dados lexicais:

Wordnet	http://wordnet.princeton.edu/
Wordnet Português	http://www.clul.ul.pt/wn/

Glossários e dicionários do domínio - PT

Dicionário da Construção Civil	http://www.ecivilnet.com/dicionario/
Construção Civil -Glossário de termos técnicos	www.forma-te.com/
Dicionário abc da construção	http://luispatricio.planetaclix.pt/abconstrucao.htm

Nota: Encontrará, anexos, alguns glossários em formato .pdf

OUTROS RECURSOS:

Motores de pesquisa de traduções:

LazyTerm	http://terminotrad.info/RogerMcKeon/LazyTerm/TerminoParesse-Rev.3e.html
Global Glossary	http://en.globalglossary.org/Default.aspx

Recursos específicos do domínio

Patorreb	http://www.patorreb.com/pt/
----------	---

Páginas web localizadas (*páginas localizadas em diferentes línguas na área da comercialização de produtos de construção*)

SIKA	http://prt.sika.com/pt/solutions_products/02.html
Grupopuma	http://www.grupopuma.com/pt/

LISTA DE MAPAS CONCEPTUAIS A LOCALIZAR

Poderá usar o espaço **Observações** para colocar dúvidas e fazer reflexões sobre:

(1) sobre o *processo de conceptualização*

- a. Ex.: apontar necessidades de alteração na representação conceptual,
- b. Ex.: apontar necessidade de alteração da relação conceptual.

(2) o *processo de localização*, nomeadamente sobre os problemas encontrados ao nível da identificação e validação dos equivalentes, como por exemplo:

- a. Inexistência de equivalentes exactos para o conceito/termo,
- b. Existência de mais do que um equivalente,
- c. Inexistência do conceito na sua língua de trabalho (na sua cultura),
- d. Inexistência/má qualidade das traduções propostas pelas aplicações de tradução.

Nota: os exemplos 1 e 2 são apenas indicativos.

	MAPAS	OBSERVAÇÕES
Global frame of	1 - global frame of the	1

the high level ontology	<i>high level ontology</i>	
Construction Process	<i>2- work process</i>	2
	<i>3- construction process</i>	3
	<i>4- project stages</i>	4
Construction Resource	<i>5- construction aid</i>	5
	<i>6- construction resource</i>	6
	<i>7- material</i>	7
Construction Result	<i>8- construction entity</i>	8
	<i>9- construction result</i>	9
Technical Topics	<i>10- technical Topic</i>	10
	<i>11- intervention programme</i>	11
	<i>12- planning system</i>	12

Apêndice C: Guide for the expert

GUIDE FOR THE EXPERT

Objectives for the localization of the conceptual maps

It is our view that in the context of a multilingual collaborative network, conceptual representation of domain knowledge is achieved through the structuring which results from negotiation and consensus obtained among the experts of the domain.

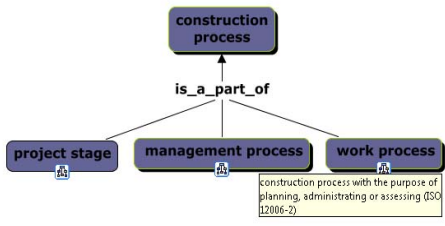
Having this into account, this process is expected, through your participation to:

3. Help develop the conceptual representations of the domain's knowledge in your own working language,
4. Promote the participation of the various experts in the shared construction of the ontology, to better understand and represent the specific requirements of their language and culture
5. Identify experts doubts and difficulties in the process

Conceptual maps analysis and localization

CONCEPTUALIZATION AND LOCALIZATION SPACE:

- Each map sent to you has a set of elements to support both conceptualization and localization. In each map you will find:

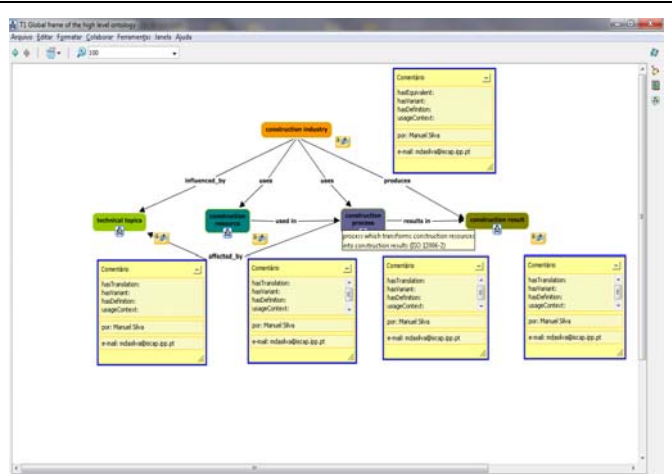
<p><i>A domain definition in English:</i> They have been posted as comments (cmap functionality) and you can access them by stopping the mouse on the concepts</p> <p><i>The definitions indicate the source of each definition. (defined in H-Know project)</i></p> <p><i>(You may had others if you wish)</i></p>	 <pre>graph TD; A[construction process] -- is_a_part_of --> B[project stage]; A -- is_a_part_of --> C[management process]; A -- is_a_part_of --> D[work process];</pre>
---	--

By using the Cmaps comments functionality (as shown in the figure) you are able to insert your suggestions/revisions and insert information like::

- d) one or more language variants
- e) definitions (in your own language)
- f) context of use (ex.: extract from a specialized text)

In each commentary box you will find the following elements that you may fill:

- 5. *hasTranslation:*
- 6. *hasVariant:*
- 7. *hasDefinition:*
- 8. *usageContext:*



As an expert you may insert, along with the equivalents, definitions or notes in this commentary area to support/reinforce your choice / present doubts.

RESOURCES AND TOOLS TO SUPPORT THE LOCALIZATION PROCESS

Domain Sources

(These sources were used as the basis for the initial conceptualization process. The definitions displayed in the conceptual maps were extracted from these sources)

UNICLASS - Unified Classification for the Construction Industry	http://en.wikipedia.org/wiki/Uniclass
Thesaurus – European Heritage Network	http://www.european-heritage.net/sdx/herein/thesaurus/consult.xsp?class=1&nr=361
OmniClass Construction Classification System	http://www.omniclass.ca/
EPIC – European Product Information Co-Operation	In annex
ISO 12006-2	In annex

TOOLS TO SUPPORT CONCEPTUAL REPRESENTATION

*The conceptual maps are in Cmap format. Please use **CmapTools** for their edition.*

IHMC CmapTools	http://cmap.ihmc.us/download/
----------------	---

TOOLS TO SUPPORT LOCALIZATION (TRANSLATION)

Web applications for **Machine translation**

Google translator	http://translate.google.pt/
Bing Translator	http://www.microsofttranslator.com/
Babel Fish	http://babelfish.yahoo.com/
Linguee (<i>corpus of exemples of translations</i>)	http://www.linguee.en/

Terminological Databases:

IATE (Terminological Database of the European Community)	http://iate.europa.eu
--	---

Lexical Databases:

Wordnet	http://wordnet.princeton.edu/
Wordnet for Italian, Portuguese, French, English and Spanish	http://multiwordnet.fbk.eu/online/multiwordnet.php

Glossaries and dictionaries of the domain - DE

Fachwörterverzeichnis für den statisch-konstruktiven Bereich des Bauwesens	In annex
Glossary Renovation Restoration	In annex

NOTE: You may had other glossaries and dictionaries

OTHER RESOURCES:

Translation research engines:

LazyTerm	http://terminotrad.info/RogerMcKeon/LazyTerm/TerminoParesse-Rev.3e.html
----------	---

Global Glossary	http://en.globalglossary.org/Default.aspx
-----------------	---

Localized (translated) webpages (*webpages localized in several languages in the area of construction products, etc.*)

SIKA	http://prt.sika.com/pt/solutions_products/02.html
Grupopuma	http://www.grupopuma.com/

LIST OF CONCEPTUAL MAPS TO BE LOCALIZED

In order to better express your views, you may use the next chart and the area **Observations**. You may use this area to present your opinions, doubts and suggestions on:

(1) the *conceptualization process*

- a. Ex.: point the need for modifications in the conceptual structure/representation
- b. Ex.: point the need for modifications in the conceptual relations

(2) the *localization process*, namely on the problems you have found when identifying and validating the language equivalents, as, for instance::

- a. Inexistence of exact equivalents for the concept/term
- b. Existence of more than one equivalent
- c. Inexistence of the concept in your working language (culture)
- d. Inexistence/poor quality of the available translations (either found in dictionaries or glossaries or obtained through a machine translation engine)

Note: examples 1 and 2 are merely an indication.

	MAPS	OBSERVATIONS
Global frame of the high level ontology	<i>1 - global frame of the high level ontology</i>	<i>1</i>

Construction Process	<i>2- work process</i>	2
	<i>3- construction process</i>	3
	<i>4- project stages</i>	4
Construction Resource	<i>5- construction aid</i>	5
	<i>6- construction resource</i>	6
	<i>7- material</i>	7
Construction Result	<i>8- construction entity</i>	8
	<i>9- construction result</i>	9
Technical Topics	<i>10- technical Topic</i>	10
	<i>11- intervention programme</i>	11
	<i>12- planning system</i>	12

Apêndice D: Entrevista aos participantes no caso de estudo

No contexto de uma rede colaborativa multilingue e multicultural, a representação conceptual do conhecimento de um domínio advém de uma estruturação obtida através de negociação e consenso entre os especialistas desse domínio, num processo onde as culturas de cada actor da rede se sobrepõem, num espaço de conceptualização mediado por uma intercultura profissional.

Tendo por base o projecto H-Know e no âmbito do estudo de Doutoramento em Linguística que estamos a desenvolver, pretendemos, com esta entrevista, analisar a implementação de uma metodologia de localização (tradução) e o seu contributo para o processo de partilha, conceptualização e representação do conhecimento do domínio da reabilitação, no contexto de uma rede colaborativas multilingue e multicultural.

Objectivos:

1. Analisar os resultados da aplicação da metodologia de localização de ontologias, através da:
 1. Análise das necessidades e dificuldades especialistas sobre o processo desenvolvido e a sua participação.
 2. Análise da opinião dos especialistas quanto ao contributo da metodologia proposta para a representação do conhecimento do domínio em termos multilingues, para a partilha de conhecimento e obtenção de consenso e, potencialmente, para a qualidade final da ontologia.

I

1) Identificação (Perfil)

Nome:

País de origem:

Área de especialidade:

Tipo de instituição:

PME/Associação

Instituto de investigação

II

**Breve caracterização do especialista quanto ao uso de recursos linguísticos e de localização
(tradução)**

*(Estas questões destinam-se a perceber a maior ou menor facilidade para participar
activamente no processo de localização)*

2. Já alguma vez traduziu?
 - a. Sim
 - b. Não

3. O recurso a traduções, no seu contexto profissional, é:
 - a. Muito frequente
 - b. Pouco frequente
 - c. Quase inexistente

4. Que recursos de tradução utiliza/utilizou?
 - a. Bases de dados terminológicas (ex.: IATE)
 - b. Bases de dados lexicais (ex.: Wordnet)
 - c. Glossários (específicos do seu domínio)
 - d. Dicionários (gerais)
 - e. Dicionários (específicos)
 - f. Tesouros
 - a. Tradução automática (ex.: Google translator)

5. Como descreveria o grau de dificuldade de uso das ferramentas de tradução existentes:
 - d. Alto
 - e. Médio
 - f. Baixo
 1. Considera que, no contexto de construção de uma ontologia, e como especialista do domínio, pode desempenhar um papel activo no encontro e definição dos equivalentes para a sua língua?
 - a. Sim
 - b. Não
 - i. (indique razões)

6. As ferramentas de localização (tradução) que utiliza/utilizou, no seu estado actual, produzem resultados com qualidade suficiente para auxiliar um especialista sem formação em tradução uma participação eficiente no processo de localização de uma ontologia?
 - c. Sim
 - d. Não

- i. (indique razões)
7. Como avalia a qualidade dos resultados obtidos com o uso dos recursos de tradução que utilizou neste processo:
- a. Alta
 - b. Média
 - c. Baixa
 - d. Sem qualidade

III

Questões sobre o processo de conceptualização e localização

(As questões seguintes destinam-se a obter feedback sobre os elementos do método)

8. Na sua perspectiva, os mapas conceptuais são uma ferramenta viável para a representação do conhecimento sobre um domínio específico em mais do que uma língua.
- c. Sim
 - d. Não
9. Os passos e elementos constantes dos mapas conceptuais – definições, identificação das fontes, propostas de equivalentes - são relevantes e auxiliam à representação do conhecimento e à sua localização nesta fase de especificação semiformal?
- a. Sim
 - b. Não
 - i. (indique razões)
10. Que outros elementos consideraria relevantes para tornar mais eficiente o processo de conceptualização e localização com base em mapas conceptuais.
11. O facto de os recursos de conhecimento (ex.: classificações) propostos para a construção da ontologia possuírem já uma representação própria do conhecimento (ex.: hierárquica)
- a. afecta ou condiciona a sua representação do domínio
 - b. facilita a sua representação do domínio?
 - i. (indique razões)
12. Que problemas encontrou ao nível da identificação e validação dos equivalentes?

- a. Inexistência de equivalentes exactos para o conceito/termo
- b. Existência de mais do que um equivalente
- c. Inexistência do conceito na sua língua de trabalho (na sua cultura)
- d. Outros (indique quais)

(As questões seguintes destinam-se a obter feedback sobre o uso do método)

13. Considera que a presença de mais do que uma língua pode contribuir para atingir um maior nível de consistência na construção partilhada da representação conceptual do conhecimento relativo ao domínio (e da ontologia)?
- a. Sim
 - b. Não
 - i. (Indique razões)
14. Considera que o processo de localização, desenvolvido na fase de especificação semiformal do conhecimento do domínio, pode contribuir para uma maior partilha de conhecimento entre os especialistas do domínio?
- a. Sim
 - b. Não
 - i. (Indique razões)
15. Considera que o processo de localização, desenvolvido na fase de especificação semiformal do conhecimento do domínio, pode contribuir para a melhoria dessas representações e, em consequência, para um aumento da qualidade da ontologia?
- c. Sim
 - d. Não
 - i. (Indique razões)
16. É a representação conceptual que influencia as suas decisões de localização ou é a localização que potencia a sua visão da representação conceptual?

Apêndice E: Interview to the participants in the case study

In the context of a multilingual and multicultural collaborative network, the conceptual representation of a domain's knowledge is based on a representation obtained through negotiation and consensus among domain specialists, in a process where the cultures of each actor of the network overlaps in a conceptualization space mediated by a professional interculture.

Departing from H-Know Project and in the scope of a PHD in Linguistics that we are developing, it is our intention with this interview to analyze the implementation of a localization methodology and its contribute to the process of knowledge sharing, conceptualization and representation in the specialized domains, in the context of multilingual and multicultural collaborative networks.

Objectives:

1. Analyze the results of the application of the ontologies localization method, through:
 - a. Analysis of the experts' needs and difficulties in what concerns the conceptualizations and localization processes and their participation in them.
 - b. Analysis of the experts' opinion as to the contribution of the proposed method in what respects the multilingual representation of domain knowledge, the knowledge sharing and consensus building process and the contribution for the final quality of the ontology.

I

1) Profile

Name:

Country of origin:

Area of expertise:

Type of Institution:

SME/Association

RTD

II

Brief description of the specialist in what respects his use of linguistic and localization resources and tools

(The following questions aim at assessing the availability of the specialist to actively engage in the localization process)

2. Have you ever translated?
 - a. Yes
 - b. No

3. On a professional level, do you need translations:
 - a. Very frequently
 - b. Not very frequently
 - c. Rarely

4. Which translation resources do you used/use?
 - g. Terminological databases (ex.: IATE)
 - h. Lexical databases (ex.: Wordnet)
 - i. Glossaries (specific)
 - j. Dictionaries (general)
 - k. Dictionaries (specific)
 - l. Thesauri
 - m. Machine translation (ex.: Google translator)

5. How would you describe the level of difficulty in the usage of the existing translation tools
 - g. High
 - h. Medium
 - i. Low

6. Do you consider that, in the context of the construction of an ontology, and as an domain expert, you can play an active role in the search and definition of the equivalents for your language?
 - e. Yes
 - f. No
 - i. (Please describe why)

7. The localization (translation) tools that you use or have used are, in their present state of development, capable of producing results with enough quality to help a

specialist without translation training to efficiently participate in the process of localizing an ontology?

- e. Yes
- f. No
- i. (Please describe why)

8. How do you assess the quality of the results obtained through the usage of the translation resources that you have used in this process?

- e. High
- f. Medium
- g. Low
- h. Without quality

III

Questions about the conceptualization and localization processes

(The following questions aim at obtaining feedback on the method's elements)

9. From your point of view, are the conceptual maps a viable tool for representing the knowledge of a specific domain in more than one language?

- g. Yes
- h. No

10. Do you think that the steps and the elements available in the conceptual maps – definitions, identification of sources, proposals for equivalents - are relevant and helpful in the process of knowledge representation and localization in this semiformal specification stage?

- c. Yes
- d. No
- i. (Please describe why)

11. What other elements would you consider to be relevant/interesting to make the process of conceptualization and localization based on conceptual maps more efficient?

12. The fact that the recommended knowledge resources for the ontology construction (ex.: classifications) already have their own knowledge representation (ex: hierarchical)

- c. Affects or limits your representation of the domain
- d. Facilitates your representation of the

domain

ii. (Please describe why)

13. What problems did you find related to the identification and validation of the equivalents?

- e. Inexistence of exact equivalents for the concept/term
- f. Existence of more than one equivalent
- g. Inexistence of the concept in you work language (culture)
- h. Other (indicate which)

(These questions aim at obtaining feedback about the use of the method)

14. Do you think that the presence of more than one language may contribute to the achievement of a higher consistency level in the shared construction of a conceptual representation of the knowledge specific domain (and of the ontology)?

- c. Yes
- d. No
- i. (Please describe why)

15. Do you think that the localization process, developed during the semi-formal domain's knowledge representation stage may contribute to a better knowledge sharing among the specialists of the domain?

- a. Yes
- b. No
- i. (Please describe why)

16. Do you think that the localization process, developed during the semi-formal domain's knowledge representation stage, can contribute to the enhancement of those representations and, as a consequence, to improvements in the quality of the ontology?

- a. Yes
- b. No
- ii. (Please describe why)

17. Is it the conceptual representation that influences your localization decisions or is it localization that enhances your vision of the conceptual representation?

Apêndice F: Entrevista - Recolha e análise de resultados

Nota: Das cinco entrevistas, duas foram presenciais, duas foram realizadas por telefone e uma foi obtida por escrito.

I

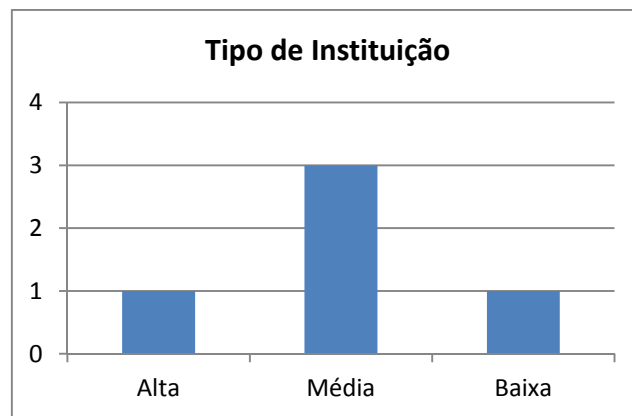
1) Identificação (Perfil)

Nome:

País de origem: *Portugal/Alemanha/França/Espanha/Sérvia*

Área de especialidade: *Engenharia Civil/Engenharia Informática*

Tipo de instituição:

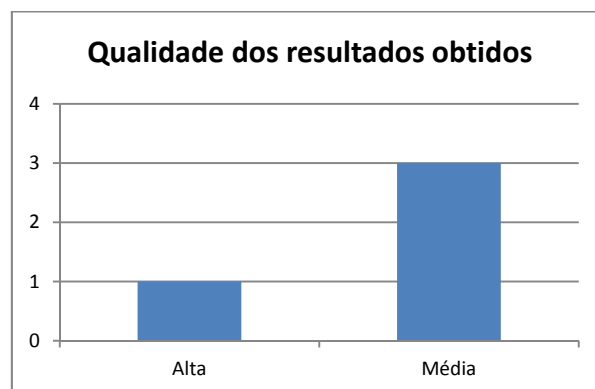


II

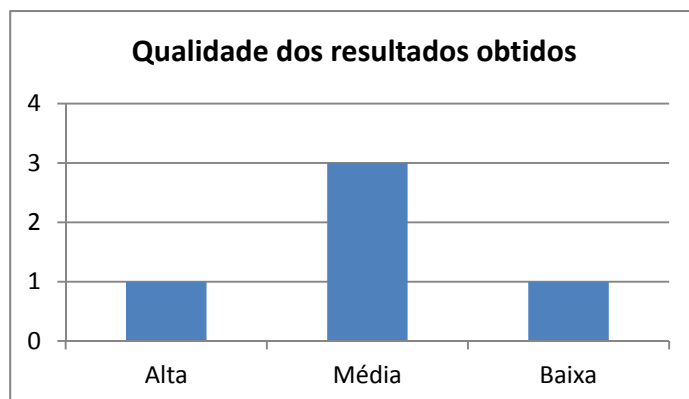
Breve caracterização do especialista quanto ao uso de recursos linguísticos e de localização (tradução)

(Estas questões destinam-se a avaliar, de modo simples, a maior ou menor facilidade para participar activamente no processo de localização)

2. Já alguma vez traduziu?



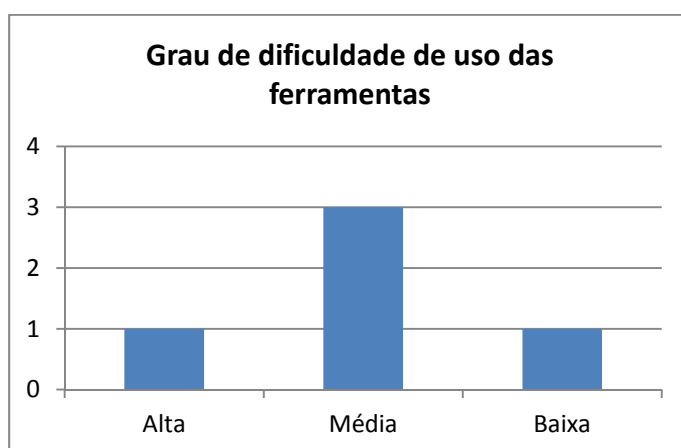
3. O recurso a traduções, no seu contexto profissional, é:



4. Que recursos de tradução utiliza/utilizou?

- | | |
|---|---|
| a. Bases de dados terminológicas (ex.: IATE) | 1 |
| b. Bases de dados lexicais (ex.: Wordnet) | 0 |
| c. Glossários (específicos do seu domínio) ¹ | 1 |
| d. Dicionários (gerais) | 3 |
| e. Dicionários (específicos) | 2 |
| f. Tesouros | 3 |
| g. Tradução automática (ex.: Google translator) | 3 |
| a. Outros: word | 1 |

5. Como descreveria o grau de dificuldade de uso das ferramentas de tradução existentes:



- i. AS_{pt}: *Considera difícil encontrar outras ferramentas para além das identificadas.*

6. Considera que, no contexto de construção de uma ontologia, e como especialista do domínio, pode desempenhar um papel activo no encontro e definição dos equivalentes para a sua língua?

i. Sim

5

j. Não

i. (indique razões)

- AS_{pt}: *formação e experiência do especialista.*
- JPM_{pt}: *O especialista tem de desempenhar esse papel. A ontologia não é um objecto puramente linguístico. O significado requer uma compreensão e uma explicação extensa. Neste processo o linguista tem um papel de mediação – suscita a discussão, desestabiliza, retira o especialista da sua zona de conforto, é provocador, promove o ambiente de negociação. Permite, com a sua presença, uma maior reflexão sobre a representação.*
- LJU_{de}: *A big experience with different languages in few technical domains helps finding the right meanings in different languages, above all in languages that have a Latin basis. To know 2 or 3 European languages is quite useful.*
- JF_{fr}: *The specialist is not an expert on ontology construction or in translation - he assumes another role - facilitates the translation and acts as an intermediary.*
- AA_{es}: *The specialists always have a holistic perspective of the domain and know the different points of view of the organizations involved in the construction of the ontology and acknowledge the need to reach a common view of the domain*

7. As ferramentas de localização (tradução) que utiliza/utilizou, no seu estado actual, produzem resultados com qualidade suficiente para auxiliar um especialista sem formação em tradução uma participação eficiente no processo de localização de uma ontologia?

a. Sim

5

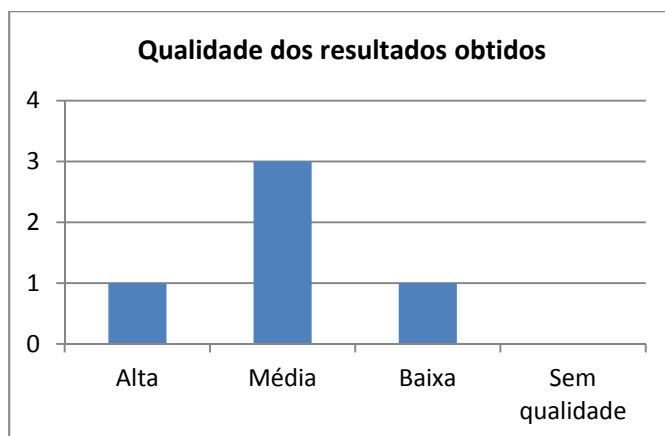
b. Não

i. (indique razões)

- AS_{pt}: *auxiliam o especialista sobretudo numa fase inicial. Permitem, em termos de produtividade, realizar ¾ do trabalho)*
- JPM_{pt}: *melhor do que não ter. São uma boa ajuda inicial.*
- LJU_{de}: *I use them just as a support tools.*
- AA_{es}: *technology is evolving and engines are more sophisticated and improving. But results depend on the features you want to use. You get better results by using specific domain tools.*

- JFr: *I never use only one tool – automatic + dictionary + other sources of information.*

8. Como avalia a qualidade dos resultados obtidos com o uso dos recursos de tradução que utilizou neste processo:



Questões sobre o processo de localização

(As questões seguintes destinam-se a obter feedback sobre a aplicação da metodologia)

9. Na sua perspectiva, os mapas conceptuais são uma ferramenta viável para a representação do conhecimento sobre um domínio específico em mais do que uma língua.

k. Sim **5**

l. Não

- AS_{pt}: *Funcionam sobretudo nesse aspecto. Permitem ao especialista interpretar a realidade representada e a tradução permite que se relacione/se ponha em perspectiva a interação entre os processos e as entre as diferentes fases (relativas ao domínio)*
- JPM_{pt}: *A língua faz parte do processo – está inserida no processo.*
- JFr: *It is a good tool and clear to see the complexity. Already in use in other departments areas / projects of the association.*

10. Os passos e elementos constantes dos mapas conceptuais – definições, identificação das fontes, propostas de equivalentes - são relevantes e auxiliam à representação do conhecimento e à sua localização (tradução)?

e. Sim

f. Não

i. (indique razões)

3

- AS_{pt}: *ajuda a definir o conceito. As palavras não possuem um só sentido. Assim delimita-se o uso da palavra e enquadra-se melhor o conceito que se está a analisar.*
- JPM_{pt}: *Mas deveria haver duas fases: Uma 1ª fase de conceptualização livre do tipo folha de papel em branco e uma 2ª fase mais restrita, guiada e com estes ou outros elementos. O processo deveria ser iterativo.*
- LIU_{de}: *they offer enough information for the basic level (semiformal), not really for the advanced one - level of higher expertise (where more relationships are needed)*
- JF_{fr}: *The main elements are good, no need for more – if the maps are too complex / have too many elements it won't be easy to work. For localization and translation that would be no problem – for the conceptualization – it would be very difficult.*
- AA_{es}: *it helps you to start the process and helps to achieve a common understanding, as you can investigate the concept through the sources and also find the translations there*

11. Que outros elementos consideraria relevantes para tornar mais eficiente o processo de conceptualização e localização com base em mapas conceptuais.

- AS_{pt}: *figuras e gráficos*
- JPM_{pt}: *Nenhum – os elementos são suficientes.*
- LIU_{de}: *Do not know enough about conceptual maps to answer appropriately to this question*
- JF_{fr}: *No need for more. No need to complexify the complexity – systematically analyzing the maps is complex enough.*
- AA_{es}: *There should be more examples of sentences involving the concept.*

12. O facto de os recursos de conhecimento (ex.: classificações) propostos para a construção da ontologia possuírem já uma representação própria do conhecimento (ex.: hierárquica)

e. afecta ou condiciona a sua percepção do domínio 1

f. facilita a sua representação do domínio? 4

iii. (indique razões)

- AS_{pt}: *sim, mas apenas em parte, dado que há casos (domínios) que não são bem retratados nessas classificações, como é o caso da reabilitação. Por outro lado há casos em que não há uma descrição correcta do encadeamento dos processos, sendo que alguns nem chegam a ser identificados.*

- JPM_{pt}: *confortável – poupa o especialista no processo de reflexão.*
- LJU_{de}: *It shortens the time to remove dilemmas.*
- JF_{fr}: *It's easier to use conceptual maps and to understand, than to use a classification like Omniclass - due to the use of elements like relations. Omniclass is like a catalog – you cannot have a global view of things - it is not possible. Maps have an advantage – the representation / the spatial division, which helps us to understand each concept.*
- AA_{es}: *It helps to think about the representation and the missing information*

13. Que problemas encontrou ao nível da identificação e validação dos equivalentes?

- | | |
|--|----------|
| i. Inexistência de equivalentes exactos para o conceito/termo | 5 |
| j. Existência de mais do que um equivalente | 5 |
| k. Inexistência do conceito na sua língua de trabalho (na sua cultura) | 3 |
| l. Outros (indique quais) | 0 |

(As questões seguintes destinam-se a obter feedback sobre o uso do método)

14. Considera que a presença de mais do que uma língua pode contribuir para atingir um maior nível de consistência na construção da representação conceptual do conhecimento relativo ao domínio (e da ontologia)?

- | | |
|---|----------|
| e. Sim | 5 |
| f. Não | |
| i. (Indique razões) | |
| • AS _{pt} : <i>Sim. Senão pode-se cair no vazio e não ter capacidade para encontrar lacunas ou ter dificuldade em identificar os elementos (conceitos) em falta)</i> | |
| • JPM _{pt} : <i>O especialista se estiver a conceptualizar pensa mais no seu contexto – concentra-se nele, recorre à sua formação e à sua experiência. Quando está em presença de mais do que uma língua é obrigado a concentrar-se também nos pontos comuns da representação – sobretudo os de mais alto nível. A presença das diferentes línguas suscita reflexões sobre os diferentes contextos (países) e coloca o especialista perante outros desafios.</i> | |
| • LJU _{de} : <i>Different languages sometimes help to get precise information on the specific term – for somebody who can understand different languages.</i> | |

- JF_{fr}: *We should always start with one language – to develop the enrichment process. Too many languages may be complicated – more than complexity you may find yourself complicating the process. We should never depart from the particular languages – we have to start with one model in one common language and adapt to the specificity of the country.*
- AA_{es}: *The existence of more than one language works as an element that leverages the common understanding of the concepts. When you have a consistent translation is a good thing for all. You assume there is consistency while talking about concepts. You achieve a common view of the different terms, which is very important to motivate participation and the use of the tool. It also helps in the communication and helps people to better know each other and the different points of view.*

15. Considera que o processo de localização, desenvolvido na fase de representação semiformal do conhecimento do domínio, pode contribuir para uma maior partilha de conhecimento entre os especialistas do domínio?

a. Sim

5

b. Não

i. (Indique razões)

- J JPM_{pt}: *Intervenção é mais substancial. No entanto o processo torna-se mais difícil e mais complexo. Devem recorrer-se a técnicas de negociação – que podem ser formalizadas. Pode haver uma metodologia para isso – para a análise sistemática dos resultados da conceptualização.*
- LJU_{de}: *See the description above – process should be interactive. You cannot find immediately the right translation.*
- JF_{fr}: *With this type of process we can see that we haven't got the same organization and we can have a better knowledge of the organization of each other, perceive the differences and have a better knowledge of what each one does and how.*
- AA_{es}: *In my opinion this is a process that should exist from the beginning, using a continuous improvement approach cycle to validate it.*

16. Considera que o processo de localização, desenvolvido na fase de representação semiformal do conhecimento do domínio, pode contribuir para a melhoria dessas representações e, em consequência, para um aumento da qualidade da ontologia?

a. Sim

5

b. Não

i. (Indique razões)

- AS_{pt}: + *granular; + discutida; + consistente.*
- LJU_{de}: *It provides different points of view by the object name creation.*
- JF_{fr}: *Because we concentrate in the quality of the map – sometime I spent one hour to translate one concept. The ontology must be exact. It is not a middle thing – this process must take to the improvement of the ontology – it considers different points of view – which together can improve the ontology.*

17. É a representação conceptual que influencia as suas decisões de localização ou é a localização que potencia a sua visão da representação conceptual?

- AS_{pt}: *É um processo interactivo. Não há uma resposta única. A representação inicial cria uma influência na tradução. Depois da tradução – é preciso fazer uma análise e interpretação para ver se representamos o que lá estava. (duplo papel do especialista).*
 - *Caracterizou, para além disso, a língua como um instrumento de medida. Na sua perspectiva, quanto maior for o conhecimento da língua, sobretudo da língua de especialidade, mais exactamente se mede, avalia, considera e interpreta a realidade. Este ponto de vista manifesta-se também no trabalho de análise conceptual e no processo de localização.*
- JPM_{pt}: *A representação conceptual surge 1ª. No entanto o especialista parte do seu ponto de vista – um ponto de vista nacional/regional/local e a representação não se pode limitar à representação do que se conhece. O especialista precisa/deve ser confrontado pelo mediador com a 2ª parte do processo para repensar e generalizar as suas ideias. De qualquer modo este trabalho neste domínio coloca a questão do especialista – quem é o especialista da reabilitação na Europa? Será que existe? Que competências tem que ter para isso??*
- LJU_{de}: *Rather the first. Does not influences – it supports the process/the analysis – both have to be together.*
- JF_{fr}: *There is an interaction – conceptual representation influences the thought. Then, during the localization process there is a sort of interaction between the term and the map. Of course you get more interference and more discussion with the presence of the maps. It is better for the brain than just a simple list – you have a vision of the whole, of the concepts and their relations.*

Apêndice G: *Corpus* do domínio da reabilitação

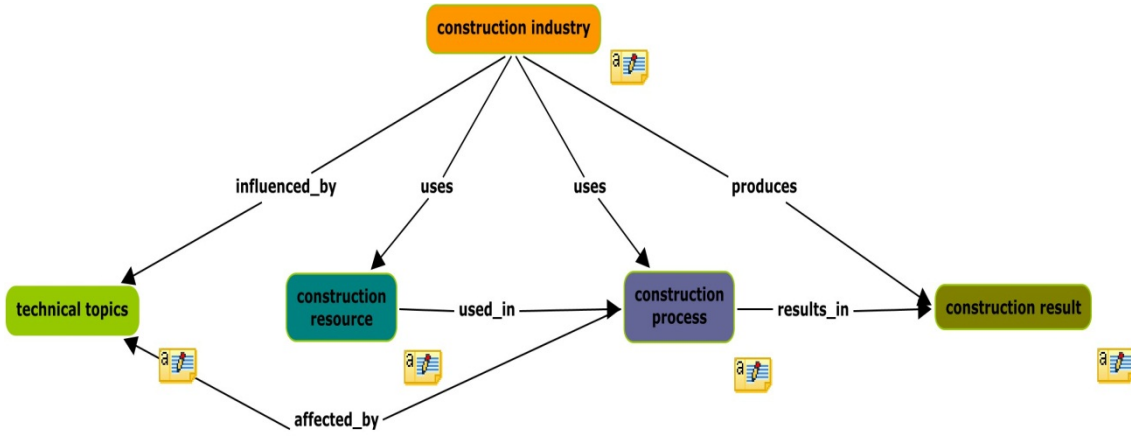
Corpus do domínio da Reabilitação

Tipo de texto	Tipo de documento	Ocorrências
<i>Texto Científico</i>	Teses mestrado Teses doutoramento	353.177
<i>Texto Técnico</i>	Monografias Artigos em conferência Revistas técnicas sobre construção Fichas sobre patologias da construção Dossiers de especialidade Revistas de especialidade	447.849
<i>Texto didático</i>	Manuais universitários Manuais de formação técnica	498.627
<i>Texto jurídico</i>	Legislação nacional e local Contrato	45 015

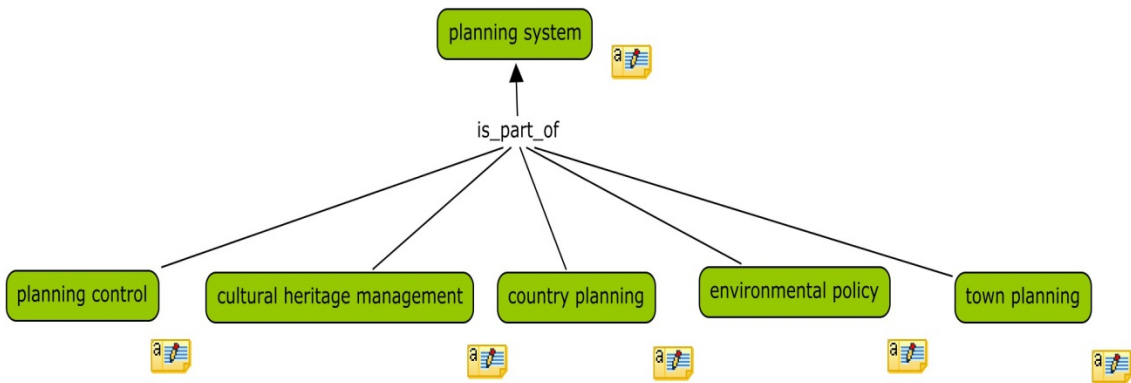
Corpus Inglês

Tipo de texto	Tipo de documento	Ocorrências
<i>Texto Científico</i>	Teses mestrado Teses doutoramento	161 027
<i>Texto Técnico</i>	Monografias Artigos em conferência Revistas técnicas sobre construção Dossiers de especialidade Revistas de especialidade	247.329
<i>Texto didático</i>	Manuais universitários Manuais de formação técnica	89.876
<i>Texto jurídico</i>	Legislação nacional e local Contrato	34.908

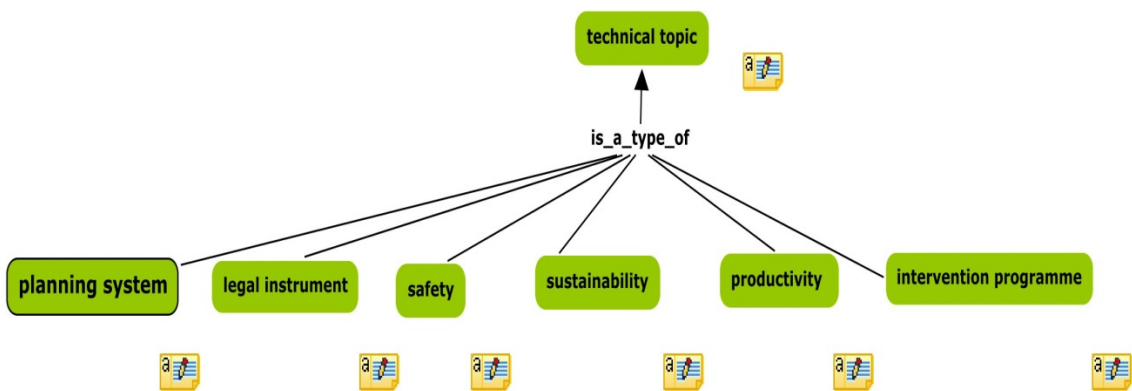
Apêndice H: Mapas conceituais: amostra



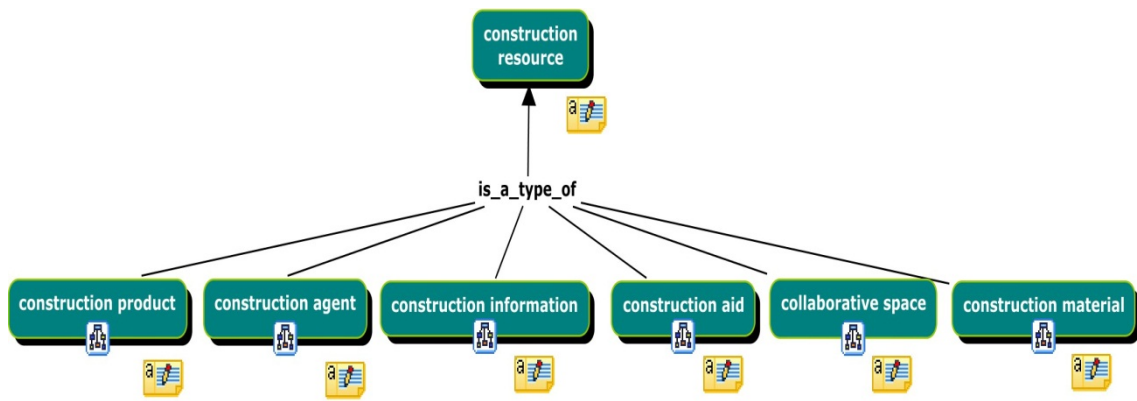
1 - Global frame of the high level ontology



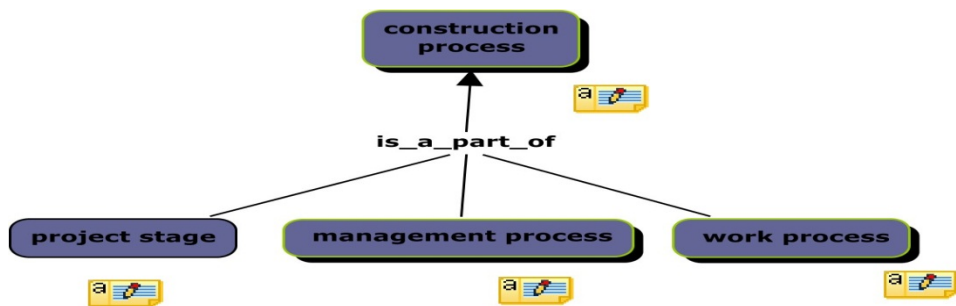
2- Planning system



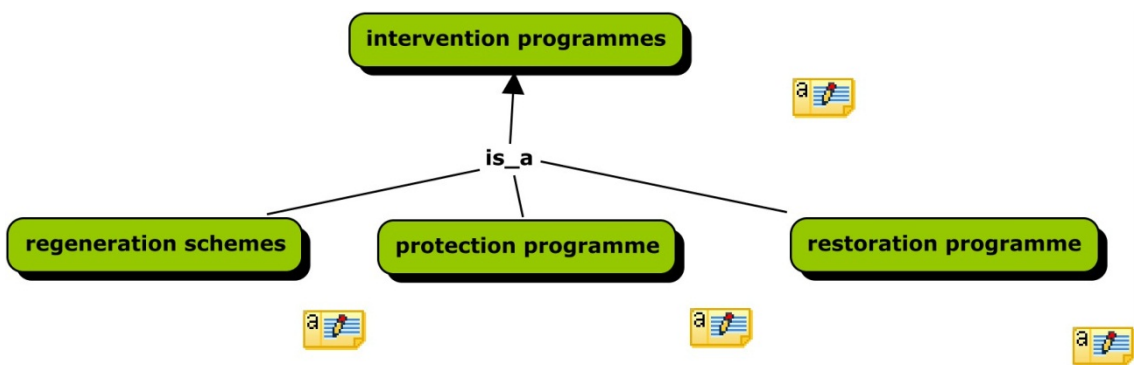
3- Technical Topic



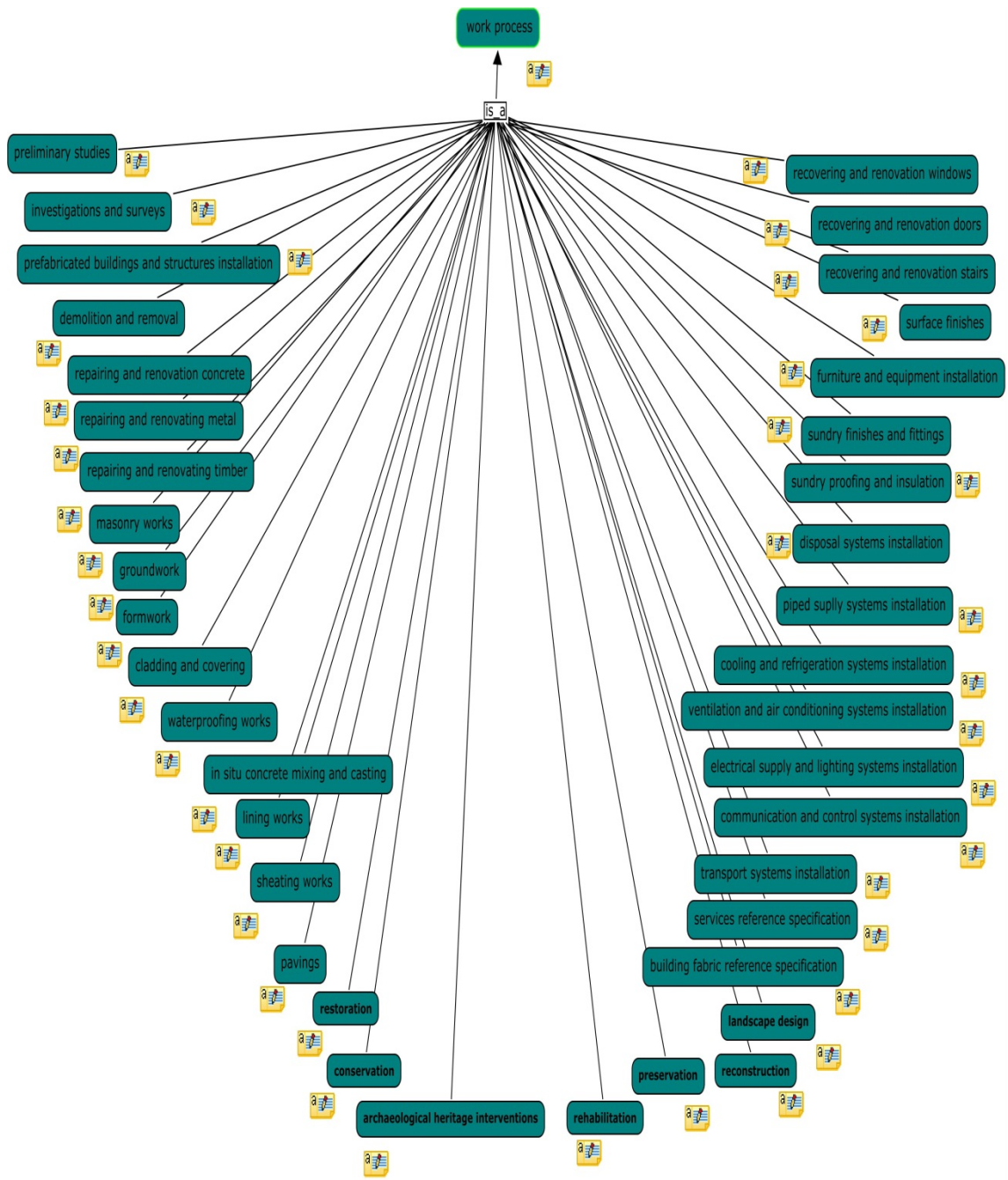
4 – Construction resource

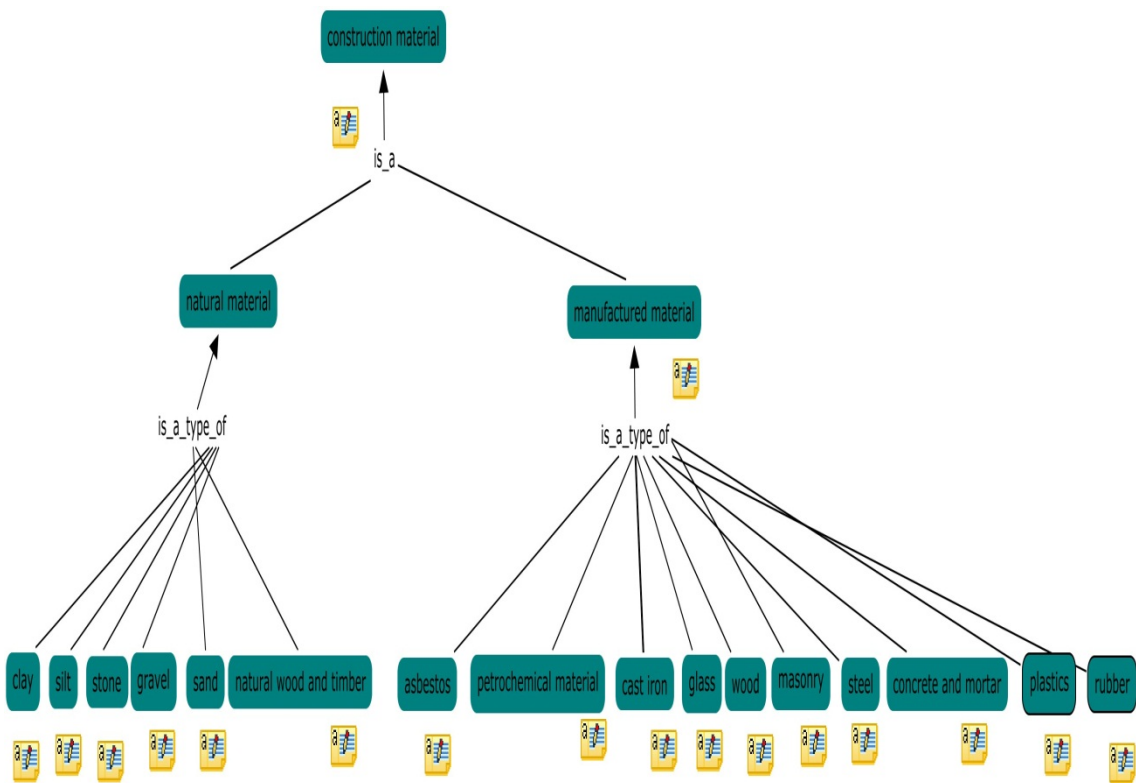
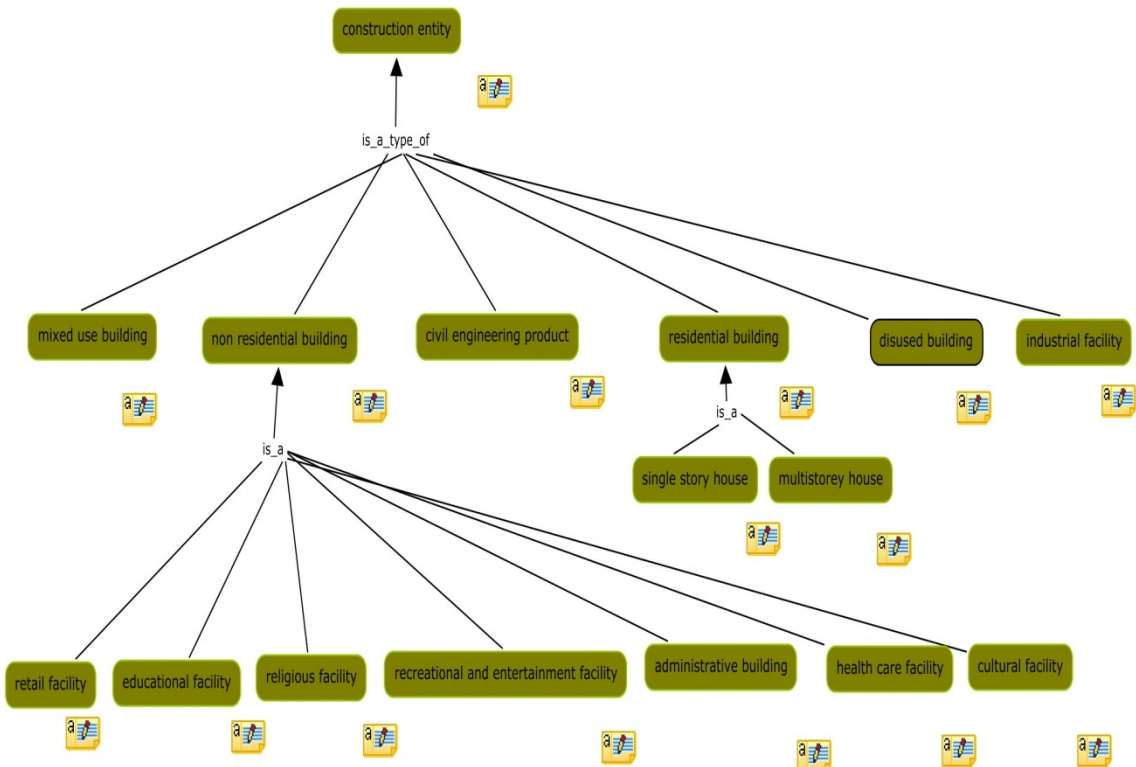


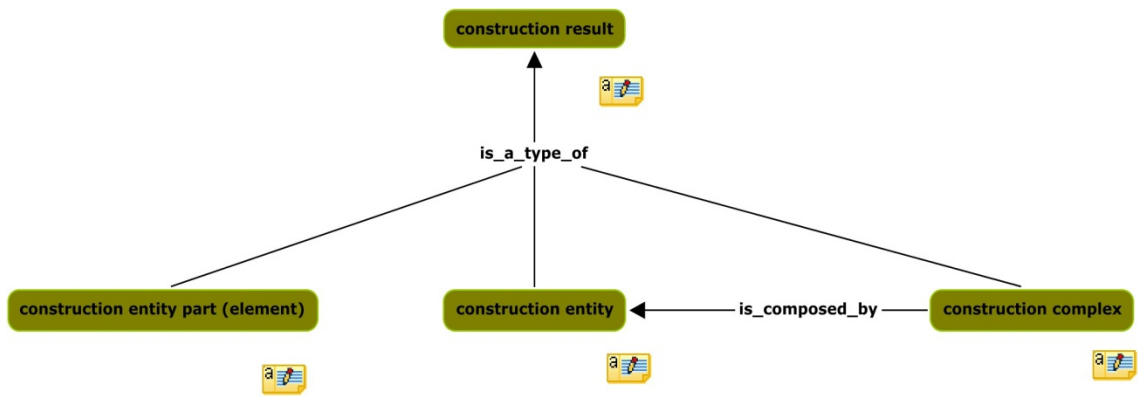
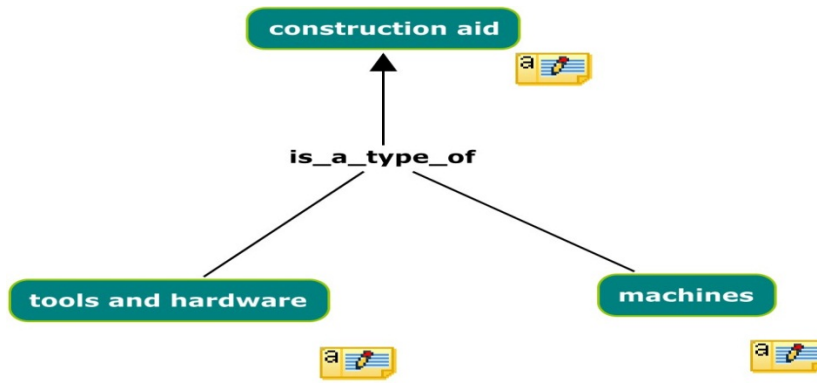
5 – Construction Process



6 – Intervention programmes







ANEXOS

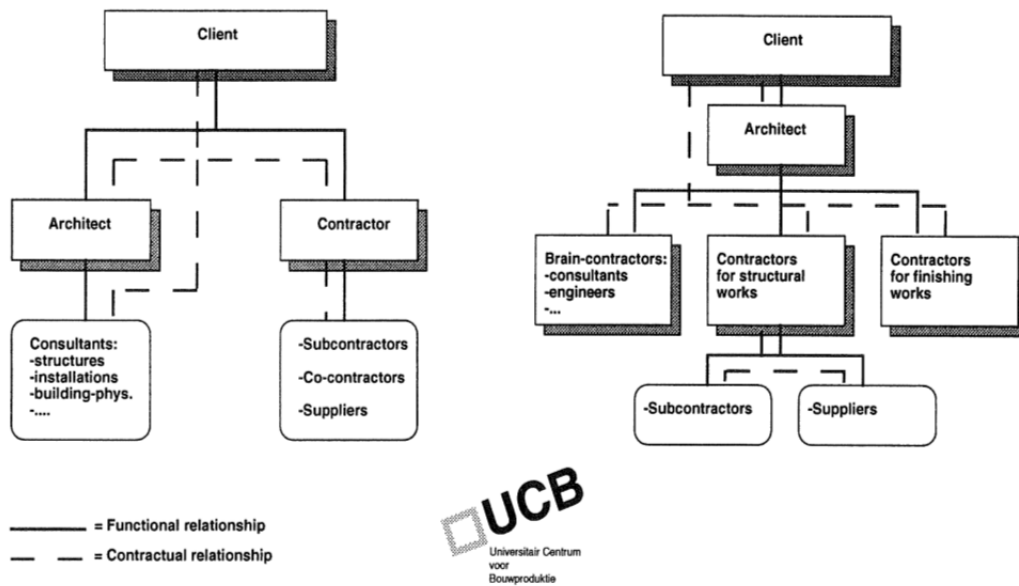
Anexo 1: Tipologias de recursos

Tipologias de classificação de recursos

- Hodge**
- *Terms Lists*
 - *Authority files*
 - *Glossaries*
 - *Dictionaries*
 - *Gazetteers*
 - *Lexicons*
 - *Classification and categories*
 - *Subject headings*
 - *Classification schemes,*
 - *Taxonomies*
 - *Categorization*
 - *Relationship lists*
 - *Thesauri*
 - *Semantic Networks*
- Gangemi et al. 1998**
- *Catalogue of normalized terms*
 - *Glossed Catalogue*
 - *Taxonomy*
- Maedche et al. 2001**
- *Text*
 - *Dictionary*
 - *Knowledge base*
 - *Relational*
- Sabou et al. 2007**
- *Unstructured*
 - *Semi-structured*
 - *Structured*

Anexo 2: Comparação entre estruturas organizacionais do processo construtivo

No exemplo seguinte comparam-se, de modo esquemático, a estrutura organizacional do processo construtivo na Holanda e na Alemanha, respectivamente, tendo o sistema alemão uma estrutura mais fragmentada que é imposta por lei².



Comparação entre a estrutura organizacional de processos construtivos (Tijhuis *et al.* 1996:129)³

² VOB *Verdingungsordnung für Bauleistungen*.

³ Tijhuis, W., Maas, G.J. (1996). *Construction process: Fragmentation or Integration? Some international experiences*. Heron, Vol. 41, N. 2, p. 129.

Anexo 3: Exemplo de Ficha Patorreb



Clarabóia – Condensações Superficiais

CONDENSAÇÕES NA FACE INTERIOR DO VIDRO DA CLARABÓIA DO ÚLTIMO PISO DE UMA HABITAÇÃO UNIFAMILIAR

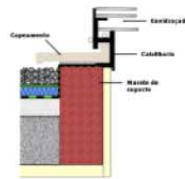
DESCRIÇÃO DA PATOLOGIA

A superfície envidraçada (clarabóia) do corredor do último piso de uma habitação unifamiliar apresentava condensações.
A habitação não dispunha de um sistema de ventilação permanente, apenas existindo dispositivos de exaustão mecânica de ar na cozinha e nas instalações sanitárias.



SONDAGENS E MEDIDAS

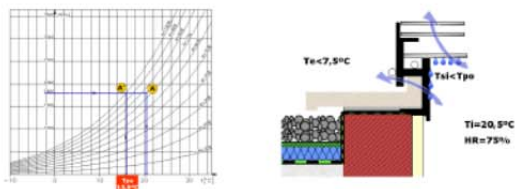
Procedeu-se à desmontagem de uma clarabóia, tendo-se observado que era constituída por caixilharia em ferro, sendo o vidro duplo (8+8+8 mm), sobreposto por uma terceira chapa de vidro, incolor.
A caixilharia era dotada de uma caleira interior para recolha de condensados, com orifícios de ventilação protegidos por tubos exteriores metálicos. Dispunha também de um dispositivo mecânico de abertura.
Os muretes em que se encontrava fixa a clarabóia eram em alvenaria de tijolo vazado com 0,15 m de espessura, sem isolamento térmico e com capeamento superior em pedra de granito.
Efectuou-se a medição da temperatura e da humidade relativa do ar e do caudal de ventilação dos compartimentos, tendo-se verificado que a ventilação era insuficiente, bem como não havia um aquecimento adequado da habitação (aquecimento intermitente).



CAUSAS DA PATOLOGIA

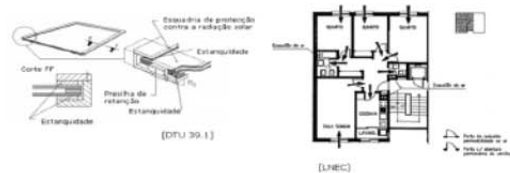
As condensações que surgiram na face interior da superfície envidraçada das clarabóias resultaram de fenómenos de condensações superficiais. Sempre que a temperatura superficial interior (T_{si}) era inferior à temperatura do ponto de orvalho (T_{po}) ocorreram condensações:
$$T_{si} = T_i - 1/h_i \cdot k \cdot (T_i - T_e) \quad (^\circ\text{C})$$

Considerando um coeficiente de transmissão térmica - K de 3 W/(m²·°C) para o envidraçado e de 6 W/(m²·°C) para a caixilharia, verifica-se que se a temperatura exterior descer abaixo dos 7,5 °C podem ocorrer condensações na superfície interior do envidraçado e quando desce abaixo dos 14°C na caixilharia. A deficiente ventilação dos compartimentos, contribuiu também para o agravamento do problema.



SOLUÇÕES POSSÍVEIS DE REPARAÇÃO

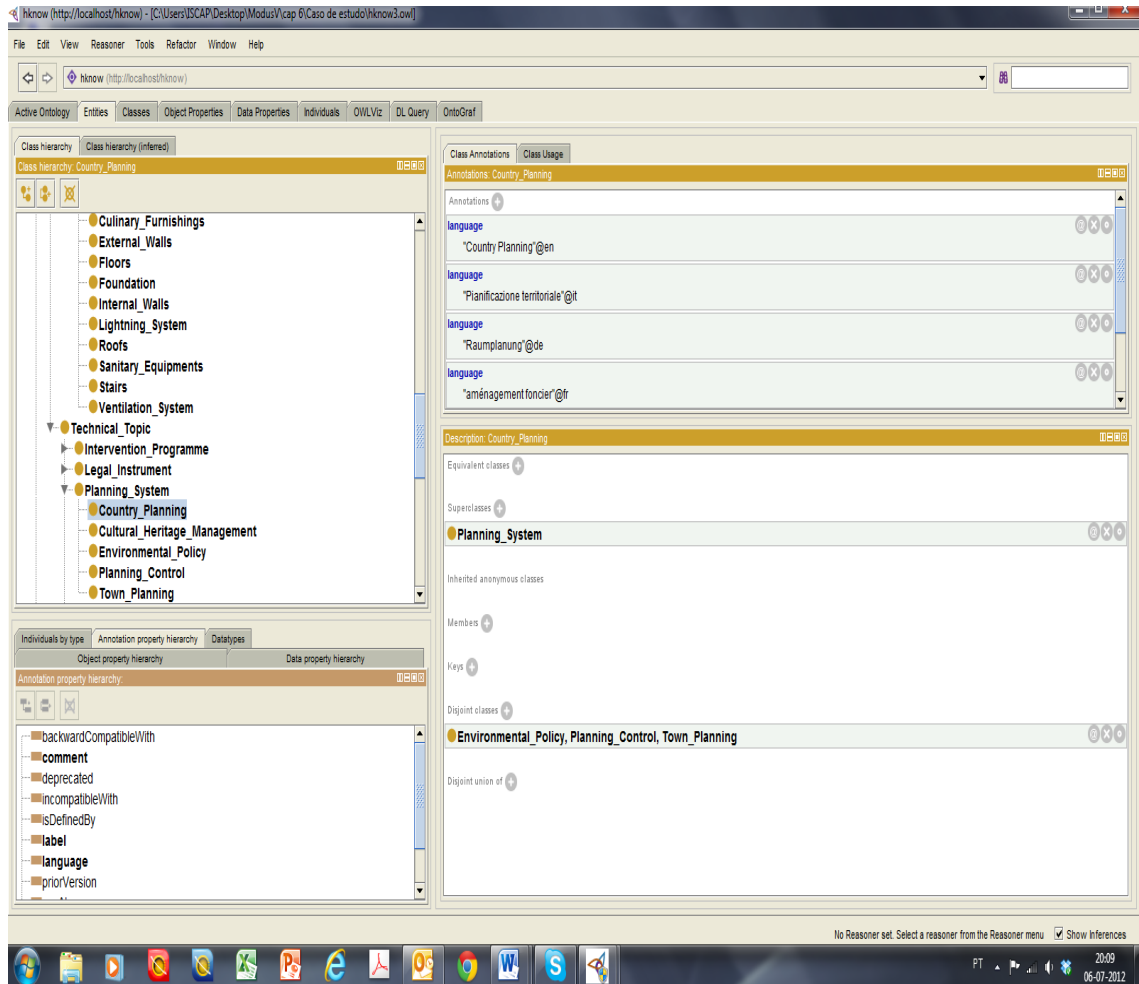
A minimização do problema das condensações na superfície interior da clarabóia passaria pela substituição do vidro, bem como da caixilharia. Para que não ocorram condensações, deveria ser aplicado um vidro duplo super-isolante, com um coeficiente de transmissão térmica - K não superior a 1,7 W/(m²·°C). A eficiência em termos de factor solar e de transmissão luminosa deveria também ser melhorada, sendo recomendável que os lanternins apresentassem um factor solar inferior a 0,20.
A substituição da caixilharia de ferro por uma caixilharia com perfis com corte térmico era também exigível.
O reforço do sistema de ventilação e o aquecimento contínuo do interior da habitação seria fundamental para melhorar o seu conforto higratérmico.



PALAVRAS-CHAVE Clarabóia, Condensações Superficiais, Vidro "Super-Isolante", Ventilação, Aquecimento

AUTORES Prof. Vasco P. de Freitas / Eng.ª Marília Sousa **REVISOR** Prof.ª Anabela Paiva

Anexo 4: Implementação da ontologia H-Know em Protégé



Implementação da ontologia H-Know em Protégé