

ARTIGO

Recebido em:
09/05/2011

Aceito em:
13/09/2012

Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação, v. 17, n. 35, p.47-74, set./dez., 2012. ISSN 1518-2924. DOI: 10.5007/1518-2924.2012v17n35p47

Aquisição de conhecimento para construção de ontologias: uma proposta de roteiro metodológico aplicado ao domínio da hematologia
Knowledge acquisition for ontologies construction: towards a methodological list of topics in the domain of hematology

Kátia Cardoso COELHO¹
Maurício Barcellos ALMEIDA²

RESUMO

Obter conhecimento especializado de um dado domínio do conhecimento é um desafio para diversos campos científicos. Em áreas correlatas às ciências da vida a terminologia é abrangente e complexa sujeita a diferentes interpretações para termos especializados. Esse tipo de problema, dentre outros relacionados, tem sido abordado há anos pela Ciência da Informação. O objetivo geral do presente trabalho é buscar alternativas para minimizar a distância entre o que o especialista tem a oferecer e aquilo que é de fato registrado como o que ele sabe. Para tal, foi proposto um roteiro para elicitación de conhecimento, para obtenção de conhecimento especializado sobre Mielopatia associada ao HTLV I. O universo empírico de dados corresponde à participação de pesquisadores do Grupo de Pesquisa em HTLV – GIPH, bem como a sua produção científica. Uma contribuição evidente da pesquisa é o roteiro metodológico descrito, que possibilitou a obtenção dos principais termos do domínio, candidatos à ontologia.

PALAVRAS-CHAVE: Aquisição de conhecimento. Expertise. Conhecimento especializado. Terminologia biomédica. Construção de ontologias.

ABSTRACT

Obtaining specialized knowledge from a specific domain is a real challenge in diverse scientific fields. Within fields related to Life Sciences particularly, the comprehensive and complex terminology may result in different interpretations for the same term. This sort of problem has



v. 17, n. 35, 2012.
p. 47-74
ISSN 1518-2924

¹ Fundação Centro de Hematologia e Hemoterapia de Minas Gerais - Universidade Federal de Minas Gerais - katiaccoelho@gmail.com

² Universidade Federal de Minas Gerais - mba@eci.ufmg.br



Esta obra está licenciada sob uma [Licença Creative Commons](https://creativecommons.org/licenses/by/4.0/)

been faced for years within the Information Science field. The general goal of this study is to propose alternatives to reduce the gap between what experts really know and what is recorded as their knowledge. In order to do this, we propose a methodology for knowledge elicitation in the specialized field of study of HTLV virus and the diseases the virus causes in human beings, namely, myelopathy associated to HTLV virus. The empirical data was obtained from researchers in the HTLV Research Group of the Hemominas Foundation (GIPH), as well as their scientific work. A noticeable contribution from this work is the proposed methodology, which eventually made possible the selection of candidate terms to be used in an ontology under construction.

KEY-WORDS: Knowledge acquisition. Expertise. Specialized knowledge. Biomedical terminology. Ontology construction

1 INTRODUÇÃO

Na Ciência da Informação, diferentes técnicas de classificação e de representação – oriundas da Biblioteconomia desde o século XIX e Documentação – têm sido utilizadas há bastante tempo: taxonomias, esquemas de classificação, vocabulários controlados, sistemas de classificação, para citar algumas. Estas técnicas representam a fundamentação básica e essencial para iniciativas de organização da informação, tanto do ponto de vista teórico quanto metodológico. No entanto, como o volume de informação a ser processado e armazenado tem aumentado significativamente nos últimos 20 ou 30 anos, representar o conhecimento e a informação para uso por pessoas e por sistemas de informação automatizados têm trazido desafios.

Nesse contexto, ontologias têm sido propostas como alternativa para criação de representações da realidade. De fato, desde os anos de 1980, tais estruturas de representação são objeto de estudo em campos como a Inteligência Artificial – como forma de apreender, formalizar e explicitar o conhecimento a ser expresso em sistemas especialistas. Na Ciência da Informação, pela tradição no uso de estruturas de categorização para organização da informação, as ontologias têm sido objeto de estudo desde os anos 90 (VICKERY, 1997). Estudos sobre ontologia em Ciência da Informação podem ser também entendidos como aqueles que buscam compreender os

processos que envolvem princípios semelhantes aos já utilizados na construção de tesouros e taxonomias.

Ontologias são também definidas a partir de um conjunto de atividades que compõem o processo de sua própria construção e compreendem: especificação, conceitualização, formalização, integração, implementação e manutenção, além das tarefas de aquisição de conhecimento, documentação e avaliação da ontologia (FERNÁNDEZ-LÓPEZ; GÓMEZ-PÉREZ; JURISTO, 1997). O estudo acerca de tais atividades é essencial para que se possa alcançar uma representação via ontologias e, portanto, de interesse científico para a Ciência da Informação (SOUZA; ALMEIDA, 2009).

O presente artigo se insere nesse contexto, enfatizando a atividade de Aquisição de Conhecimento. Trata-se de uma atividade fundamental, mas carece ainda de metodologias confiáveis para a obtenção de conhecimento, principalmente para obtenção de especialistas de domínio (HAYES-ROTH; WATERMAN; LENAT, 1983; ZAF; MCNEESE; SNYDER, 1993; MASTELA, 2005). Como alternativa, apresenta-se uma proposta de roteiro metodológico aplicado em campo no domínio da hematologia.

Para a realização da pesquisa desenvolveu-se uma metodologia baseada nas contribuições provenientes das áreas estudadas, bem como nas técnicas empregadas na atividade de AC, a partir de pessoas. O universo empírico de dados corresponde à participação de pesquisadores do Grupo de Pesquisa em HTLV - GIPH, na qual buscou elicitar o conhecimento dos especialistas no contexto da Mielopatia Associada ao HTLV (HAM/TSP), doença com a qual se estabeleceu o recorte do presente trabalho. O GIPH é um grupo de pesquisa em HTLV coordenado pela Fundação Centro de Hematologia e Hemoterapia de Minas Gerais – Hemomins e instituições parceiras.

O restante do presente trabalho está organizado como segue: a seção 2 apresenta na revisão de literatura uma visão geral sobre AC, perpassando de modo particular pela Ciência da Informação e o domínio da Biomedicina. Ambos de relevância para a pesquisa. A seção 3 apresenta o contexto e o roteiro metodológico, fundamentado na revisão de literatura. A seção 4 apresenta os

resultados e a seção 5, finalmente apresenta considerações finais e perspectivas de trabalhos futuros.

2 REVISÃO DE LITERATURA

A revisão de literatura se baseou em artigos científicos de periódicos, em trabalhos apresentados em congressos, teses e dissertações, obtidos no Portal de periódicos CAPES. Procurou-se por estudos que abordassem aquisição de conhecimento, elicitación de conhecimento, técnicas e métodos diversos. Este levantamento buscou ainda contextualizar a atividade AC, no âmbito da Ciência da Informação e da Biomedicina. Por fim, teve o propósito de situar o tema no contexto empírico e teórico.

O critério de busca se baseou em palavras-chaves simples e compostas envolvendo termos como: aquisição de conhecimento; knowledge acquisition, elicitación de conhecimento; knowledge elicitation, representação do conhecimento, knowledge representation, ontologia; ontology; ontologies; análise de domínio e domain analysis. A busca se concentrou nas áreas correlatas à Ciência da Computação/Engenharia do conhecimento, por ter sido a pioneira em estudos sobre AC; Ciência da Informação e Biomedicina por serem áreas de principal interesse para o presente trabalho. A pesquisa no “Portal de Periódicos CAPES” não se limitou a uma base de dados específica como a Web of Sciences, a escolha aconteceu a partir da seleção de áreas, como as Ciências Sociais Aplicadas.

2.1 Uma visão geral sobre Aquisição de Conhecimento

A expressão AC foi empregada na década de 1980 para se referir ao estudo da expertise das pessoas para representação em sistemas computacionais (BOOSE E GAINES, 1989; MILTON et al, 2006; MILTON, 2007). Nos anos 1990, em Gestão do Conhecimento, o mesmo tipo de atividade passou a integrar um conjunto de estratégias pioneiras que visavam apreender o conhecimento individual e convertê-lo em organizacional (CHOO, 2003). No âmbito das ontologias, a AC é uma etapa do processo de sua construção

(FERNANDEZ, GOMEZ-PEREZ e JURISTO, 1997). Em Biblioteconomia e Ciência da Informação, atividades de AC têm lugar, por exemplo, na construção de vocabulários controlados para representar informação e conhecimento em sistemas de recuperação da informação, na avaliação de sistemas de navegação da internet, bem como em ambientes de pesquisa em que é necessária a identificação e a conteúdos (NATIONAL, 2005). A AC acontece também da interação do bibliotecário com o usuário ou especialista para endosso sobre terminologia, como forma de prospectar novos termos e relacionamentos, bem como de confirmar o uso efetivo dos mesmos (LANCASTER, 1987).

Aquisição de Conhecimento (AC) é um termo amplo que compreende extrair, estruturar e organizar o conhecimento de uma ou mais fontes (TURBAN, 1992). É uma tarefa indispensável para a coleta de informações visando à organização e validação do conhecimento com a finalidade de representação (MASTELLA, 2005; MILTON, 2007).

Para as tarefas de AC, diferentes técnicas de têm sido propostas, tanto para a obtenção de conhecimento proveniente das pessoas, quanto em documentos. A AC, quando voltada para obtenção do conhecimento de pessoas é referenciada com o termo elicitación do conhecimento. De modo geral, trata-se de uma atividade composta por um conjunto de tarefas que empregam técnicas manuais e baseadas em computador (BOOSE e GAINES, 1989; DELUGACH e WOLF, 1996; BURGE, 1998; SHADBOLT, 2005; GAINES, 2003; MILTON et al 2006). Mastella (2005) considera que as técnicas de AC a partir de fontes documentadas, por ser de natureza explícita e registrada, encontram-se melhor estruturadas, embora sejam necessárias validações sistemáticas do especialista. Por outro lado, pessoas possuem diferentes visões de mundo e maneiras pessoais de solucionar problemas, tornando complexa a tarefa de AC.

Do ponto de vista teórico, as teorias e métodos que oferecem suporte à AC são obtidos em disciplinas acadêmicas diversas (PAYNE et al. 2007). As teorias que sustentam a capacidade de adquirir, representar, verificar ou validar conhecimento são provenientes de áreas como a Ciência da Computação, Ciência Cognitiva, Educação, Linguística, Semiótica e Psicologia.

Em Ciência da Computação, exemplos são os trabalhos pioneiros de Newell e Simon (1975), e de Compton e Jansen (1989). Sob essa perspectiva, a capacidade de adquirir e representar o conhecimento em um formato computável é apresentada pela hipótese de símbolos físicos (NILSSON, 2007). Essa hipótese postula que o conhecimento é constituído por símbolos que representam a realidade e pelas relações entre eles. Nesse contexto, inteligência corresponde à habilidade de manipular logicamente símbolos e relacionamentos.

Na Psicologia, as bases da atividade de AC se encontram em trabalhos seminais como a Teoria dos Construtos Pessoais proposta por Kelly (1955). Sob esse enfoque, a transferência de conhecimento constitui a base psicológica para abordar a AC: pessoas transferem sua expertise de forma que outros sejam capazes de replicar seu desempenho. Nessa mesma linha, a Ciência Cognitiva representada, por exemplo, pelo modelo Hawkins (HAWKINS, 1983), identifica a transferência de expertise como meio para a elicitación do conhecimento.

A Semiótica contribui com a perspectiva da tríade, exemplificada por Ogden-Richards (1923), e citada por Campbell (1998) como uma base teórica para AC. Tal teoria consiste do "estudo dos sinais" e de como o significado é compreendido sejam esses sinais transmitidos individualmente ou em grupo. A tríade semiótica é representada por um triângulo em que os vértices são: i) os símbolos, artefatos de representação de uma unidade de conhecimento (ex: texto ou ícones); ii) os referentes, que são objetos ou entidades; e iii) os pensamentos ou referências, unidades de conhecimento baseadas no entendimento do agente que utiliza tal conhecimento.

Na Lingüística encontram-se iniciativas em que o conhecimento é extraído por meio da análise coleções de textos narrativos. A análise de sub-linguagem é uma abordagem descrita para a AC em coleções de texto. Tal base teórica foi descrita por Zellig Harris (1976) em seu trabalho sobre a natureza do uso da linguagem em domínios altamente especializados, caracterizados pela estrutura gramatical regular e reproduzível. Esses recursos e gramáticas podem ser descobertos através da aplicação de padrões de reconhecimento manuais ou automatizados para grande corpus de linguagens de um domínio.

Cabe ainda citar que a literatura AC em biomedicina apresenta propostas específicas para representação do conhecimento médico e biológico. O detalhamento de tais propostas está além dos objetivos desse trabalho, mas aqueles interessados podem obter detalhes em Tsumoto (1998), Vita et al (2006), Payne et al (2007), Hoehndorf et al (2009), Stehr et al (2009), e Van Auken et al (2009).

Cada um dos campos apresentados tem contribuído para a formação, a compreensão e a evolução da atividade de AC e seu desdobramento em contextos distintos. Somando-se a esses estudos, é relevante apresentar como tal atividade vem sendo abordada pela Ciência da Informação.

2.2 Aquisição de conhecimento em Ciência da Informação

Na Ciência da Informação, alguns trabalhos que abordam aquisição ou elicitación de conhecimento o fazem sob diferentes enfoques e para fins diversos: Recuperação da Informação (SPINK; SOLLENBERGER, 2004); Classificação (TANGMANEE, 1994; FIDEL, 1994); construção de Tesouros (KWASNIK, 1994), Ontologias (LEENHEER, 2008); Estudo de Usuários (PISANSKI; ZUMER 2010); Gestão do Conhecimento (GARCIA-PERES, 2007; JACKSON; WEBSTER, 2007; COFFEY, 2008; PITT, 2008; BEERS, 2009; KWONG; LEE, 2009; EL HARBI 2011), Bibliotecas e Serviços de Informação (MEI-MEI, 2003); Representação do Conhecimento (PREISS, 2000; JOHN, 2003).

Spink e Sollenberger (2004) apresentam um estudo empírico de elicitación relacionado às tarefas de busca durante a Recuperação da Informação (RI). Esses autores utilizaram um corpus de transcrição de áudio relativo à interação de 30 usuários de determinada universidade, com três profissionais treinados para busca intermediária. O trabalho proposto discute modelos das tarefas relacionados a RI, na qual constataram que a análise de elicitación fornece um modelo baseado em uma melhor compreensão dos usuários finais; das pessoas que buscam informação; das necessidades de informação dos intermediários de busca e dos propósitos durante uma interação de RI.

Pisanski e Zumer (2010) apresentam um estudo relacionado à AC, sob o enfoque de modelos mentais de usuários no universo bibliográfico, no qual

verificam se os requisitos funcionais para registros bibliográficos (FRBR - functional requirements for bibliographic records) se comparam ao modelo conceitual do universo bibliográfico. Nesse estudo, Pisanski e Zumer (2010) buscavam identificar os elementos dos livros percebidos pelos usuários, além de verificar se estes usuários estavam inclinados a descrever atributos diferentes daqueles das entidades FRBR e se prestavam atenção em todos os detalhes. Para os autores, um dos maiores trunfos deste estudo é a riqueza e a variedade dos dados obtidos.

Kwong e Lee (2009) propõem a AC no contexto da gestão do conhecimento, no qual propõem elicitare o conhecimento sobre gerenciamento de confiabilidade a partir de especialistas. Para tal, desenvolveram um método para elicitare conhecimento tácito dos engenheiros de confiabilidade, por meio de narrativas e mapeamento cognitivo. Anotações foram empregadas alternativamente para registrar as transcrições das entrevistas ou dos outros dados documentais. Para os autores, trabalho forneceu um exemplo real para apoiar as fases de aprendizado do indivíduo e de um grupo no nível organizacional.

Também no contexto da gestão de conhecimento, Jackson e Webster (2007) apresentam um trabalho que enfoca elicitare de conhecimento dos formuladores de políticas de uma agência do governo australiano, para a criação de um inventário dos tipos de conhecimento necessários para se desenvolver políticas complexas para o gerenciamento dos recursos naturais. De acordo com os autores, este inventário foi usado para investigar soluções de gestão de conhecimento para lidar com a perda de informações, desgaste de pessoal e eficácia operacional. A metodologia proposta buscou combinar metodologia de sistemas flexíveis, mapeamento cognitivo causal e brainstorming para fornecer um conjunto de requisitos de conhecimento. Segundo os autores, o componente substancial da pesquisa foi o método de elicitare do conhecimento e da ontologia resultante, que contém elementos de contextualização social e raciocínio para a ação organizacional e de decisões.

Na perspectiva das ontologias, o trabalho de Almeida (2006) apresenta a atividade de AC como parte do processo de construção de ontologias, em que

são empregadas técnicas como brainstorming, entrevistas não estruturadas e estruturadas, análise informal e formal de documentos. No referido trabalho, tais técnicas tinham as seguintes funções, a saber: i) entrevistas não estruturadas que tinham por objetivo obter termos candidatos a conceitos do modelo para a memória organizacional; ii) análise de documentos, em que foi empregada a técnica de Análise de Assunto, para análise dos documentos como documentos internos, documentos externos, documentos auxiliares, registros, anotações, entre outros, manipulados pelos funcionários para desempenho de suas atividades; iii) análise de cenários; iv) técnica JAD - Joint Application Design, com vistas a obter consenso e interatividade, uma vez que tal técnica possibilita discussões em grupo, onde as decisões são baseadas em consenso; v) análise de cenários, pela facilidade de obtenção de dados sobre as necessidades e atividades dos funcionários das empresas, através de narrativas (ALMEIDA, 2006).

2.3 Aquisição de conhecimento na construção de ontologias

A atividade de AC como parte de um conjunto de atividades que compõem o próprio processo de construção da ontologia é abordado por Fernandez, Gomez-Perez e Juristo (1997) que a organiza em um conjunto de sub-atividades: especificação, conceitualização, formalização, integração, implementação e manutenção, além das tarefas de aquisição de conhecimento, documentação e avaliação da ontologia. Embora a atividade de AC para obtenção de conhecimento de um dado domínio, ocorra simultaneamente às demais, notou-se que tal atividade ainda se encontra descrita de forma superficial. Não foi encontrado na literatura consultada, detalhamentos a respeito desta etapa, mesmo que referenciada como sendo essencial e relevante para o processo de construção de ontologias (FERNÁNDEZ-LÓPEZ; GÓMEZ-PÉREZ; JURISTO, 1997; FERNÁNDEZ-LÓPEZ, 1999; FERNÁNDEZ-LÓPEZ; GÓMEZ-PÉREZ, 2002; CORCHO et. al, 2003). Considera-se, dessa forma, que aprofundar na temática da AC, a fim de buscar melhorias, bem como identificar problemas decorrentes desta atividade, é caminho a seguir para a devida representação do conhecimento, via ontologias, de um dado domínio.

Na busca de tal ação, partiu-se da premissa que considera que ontologias são construídas visando um vocabulário único, baseado em termos definidos a partir do consenso de membros de determinada comunidade (ALMEIDA; BAX, 2003). Assim, é relevante para os propósitos do presente trabalho, que se identifique o domínio do conhecimento e a comunidade que o constitui – considerando tal comunidade aquela com respaldo para falar sobre o domínio, de modo a obter desta comunidade, os termos relevantes acerca do domínio. Tal respaldo é conferido pelo caráter científico da comunidade, uma vez que, ontologias bem fundamentadas não são construídas a partir do senso-comum, mas por membros de um campo científico (SMITH, 2003).

3 UMA PROPOSTA DE ROTEIRO METODOLÓGICO

Para as atividades de AC no campo da Biomedicina propõe-se um roteiro metodológico em que o conhecimento obtido do especialista é obtido, registrado e validado sistematicamente. A seção 3.1 descreve pesquisa em andamento e o contexto em que se desenvolvem atividades de AC. A seção 3.2 descreve o roteiro para AC em Biomedicina, a partir de pesquisa em andamento no campo da hematologia.

3.1. Pesquisa em andamento

O presente trabalho integra uma pesquisa em andamento no âmbito do BloodProject³, uma iniciativa para organização do conhecimento no domínio da hematologia e hemoterapia. Os especialistas que participaram das atividades de AC são membros do Grupo Interdisciplinar de Pesquisas em HTLV (GIPH). O GIPH é coordenado pela Fundação Hemominas, que desde 1997 estuda o vírus HTLV (vírus linfotrópico de células T humanas) e acompanha indivíduos infectados. Este grupo é formado por pesquisadores desta instituição, em colaboração com uma universidade federal, um hospital e um centro de pesquisa de abrangência nacional.

³ Disponível na Internet em <<http://mba.eci.ufmg.br/BLO>>. Acesso: 22 outubro 2010.

A coleta de dados foi feita em fontes primárias (os próprios especialistas) e em fontes secundárias (literatura especializada, principalmente publicações do grupo). Um roteiro foi proposto, como forma de realizar a pesquisa fim, ou seja, buscar melhores práticas na atividade de AC em domínios correlatos à biomedicina. O roteiro, bem como as etapas de sua elaboração é apresentado na seção 3.2 como contribuição do presente trabalho.

3.2. Técnicas e instrumentos para coleta de dados

A seleção das técnicas para coleta de dados se baseou nas propostas de AC citadas na revisão de literatura. Considerou-se a AC definida em termos de suas etapas, sendo que as tarefas foram combinadas conforme cada etapa do processo. A seleção das técnicas se baseou nos seguintes critérios: i) possibilitar a compreensão do domínio do conhecimento; ii) identificar expertise dos especialistas do domínio; iii) permitir a interação com especialistas de modo individual; iv) viabilizar a elicitación do conhecimento do domínio; v) orientar a organização do conhecimento elicitado. Vale ressaltar que não foram consideradas técnicas que se propunham elicitar conhecimento, cujos objetivos fossem diferentes da proposta do presente trabalho, como, por exemplo, de verificar como determinado especialista resolveria determinados problemas, em busca de possíveis soluções etc. Tal ação seria necessária se o contexto fosse o desenvolvimento de sistemas, como preconizado pela engenharia do conhecimento, o que não é o propósito desta pesquisa.

Dentre as técnicas pesquisadas avaliaram-se aquelas consideradas adequadas ao contexto da presente pesquisa. Por exemplo, verificou-se que o grupo escolhido para as sessões de elicitación de conhecimento possuía grande número de publicações científicas. Desse modo, optou-se por métodos que viabilizassem o mapeamento do domínio a partir da produção científica. Nesse caso, a análise de domínio foi considerada como um método adequado, por possibilitar ao pesquisador identificar o domínio por meio dos textos científicos produzidos pelos especialistas. Os métodos selecionados para a presente pesquisa são apresentados à seguir, de acordo com as etapas definidas: i)

Entrevista, por ser uma técnica que possibilita acesso de modo formal ou informal ao especialista, permite obter uma visão geral do domínio, e elicitava o conhecimento do especialista de modo verbalizado; ii) Imersão na Literatura, possibilita obter os conceitos básicos do domínio em questão, para conhecimento teórico mínimo requerido; iii) Análise de Domínio, pela visão acerca do domínio, em que é possível elencar os principais elementos que o constituem. Em especial pela possibilidade de aplicar métodos como a bibliometria e análise de citação, métodos válidos e amplamente utilizados na Ciência da Informação e Biblioteconomia, para mapear campos do conhecimento a partir do domínio; iv) Análise de protocolo, por ser um método que auxilia na transcrição de entrevistas, que objetiva analisar os registros gerados pelos especialistas; v) Teachback, por ser uma técnica que visa obter um feedback dos especialistas, a partir da descrição de uma parte do conhecimento adquirido em sessões anteriores. Além disso, por auxiliar na correção de equívocos conceituais e ajuda a esclarecer mal entendidos.

Os instrumentos para a coleta de dados foram desenvolvidos ou adaptados conforme a necessidade das etapas. Entrevistas não estruturadas, entrevistas estruturadas e entrevistas de feedback fizeram uso de três formatos: i) entrevistas com perguntas abertas; ii) entrevistas para elicitava o conhecimento acerca da doença; iii) entrevista de feedback não estruturada. Além disso, foram utilizados: formulários, como levantamento de especialistas, roteiros de entrevista, agenda para sessão de aquisição de conhecimento, síntese de entrevista; template do *Protegé Frames* com roteiro para entrevista estruturada, a partir do raciocínio médico criado por Scheuermann (2009); e gravação: para registros dos contatos e transcrição.

Para entendimentos acerca do roteiro da entrevista de elicitava o conhecimento, cabe esclarecer conceitos utilizados como referência. Os termos descritos, compõem o “quadro terminológico”, que abrange as doenças, suas causas, manifestações, diagnóstico e outras entidades relacionadas ao modo como as doenças são reconhecidas e interpretadas pela medicina (SCHEUERMANN et al., 2009).

Na fase chamada processo etiológico, considera-se que há um corpo humano saudável, com características normais de acordo com parâmetros médicos. Como resultado desse processo, ocorre uma mudança física no indivíduo, dando origem à doença. Na manifestação pré-clínica da doença, o corpo desenvolve desordens, que são portadoras de disposições. Tais disposições são naturalmente associadas à existência de entidades, por exemplo, a disposição do corpo humano para ficar doente e da disposição da fruta amadurecer (SMITH, 2008). O paciente então, em alguns casos, já percebe mudanças no organismo, embora ainda não existam sinais ou sintomas.

A fase curso da doença começa com a manifestação da doença. Nesse momento, a desordem se manifesta por meio de sintomas e sinais (manifestações clínicas), sendo o primeiro aqueles que os pacientes conseguem sentir ou identificar e o segundo, que é determinado pelo médico através de exames físicos e testes laboratoriais. Nesta fase, é possível determinar os fenótipos clínicos – totalidade de sinais e sintomas-, que é a principal característica observável da doença.

Na fase resposta terapêutica, uma amostra do paciente é obtida de alguma parte do corpo, a fim de submetê-la a testes laboratoriais. Esses resultados são registrados no prontuário médico como um quadro clínico. O quadro clínico é interpretado pelo médico na busca pelo diagnóstico. Nesse ponto, é possível que se estabeleça um plano de tratamento, baseado neste diagnóstico, de modo que o corpo possa retornar à normalidade. O plano é o resultado do diagnóstico fundamentado no processo de interpretação do quadro clínico. O quadro clínico é composto por registros da representação de sintomas e sinais, bem como os resultados do exame físico e laboratorial.

A adoção de tal perspectiva para o roteiro da entrevista decorre-se do fato de se tratar de uma abordagem cujas definições são apresentadas como referência para melhores práticas no desenvolvimento de ontologias, promulgadas no OBO Foundry, conforme defendido por Smith et al (2008) e Scheuermann et al (2009).

O desenvolvimento deste percurso da doença foi adaptado na template do Protegé-Frames e apresentado a um especialista que não faria parte das sessões de elicitação, para considerações da organização e entendimentos acerca dos conceitos utilizados, conforme proposto pelos autores de referência. O exemplo apresentado na figura 1 mostra um fragmento do roteiro final.

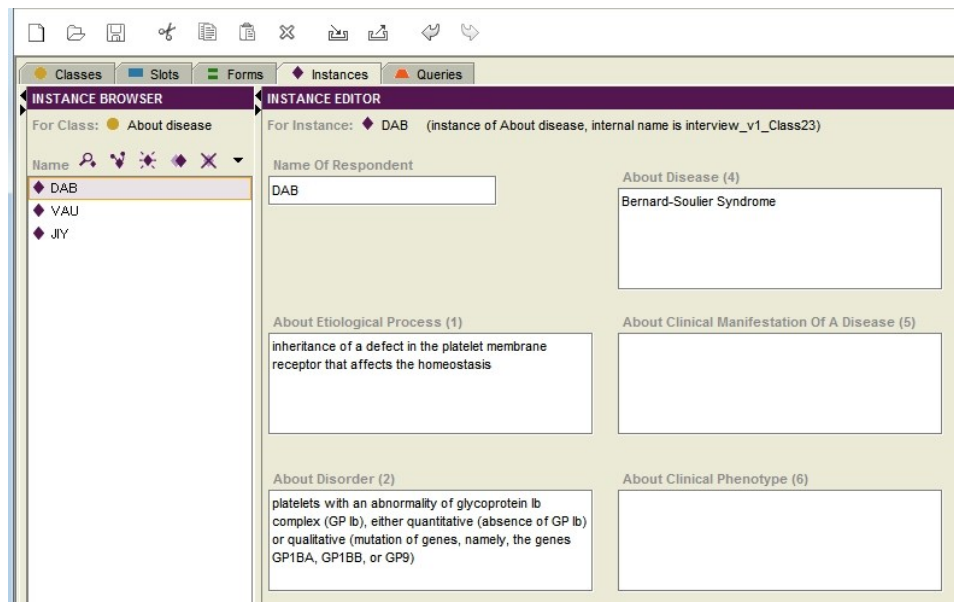


Figura 1: Roteiro no Protegé-Frames

FONTE: Desenvolvido pelos autores

3.2 Proposta de roteiro para AC

Para maior clareza, dividiu-se a coleta de dados de acordo com a origem dos dados: dados primários e dados secundários. Os dados secundários correspondem à (1) Fase de Levantamentos para o planejamento da atividade de AC. Nessa fase, os dados obtidos não exigiram em primeira instância contato com os especialistas para fins de elicitação de conhecimento. Por sua vez, os dados primários, por se tratarem de atividades inerentes à elicitação de conhecimento, exigem contato direto com especialistas. Os dados primários correspondem às fases denominadas respectivamente (2) Contato (3) Validação. O quadro 1 apresenta uma visão geral do método proposto. Cabe

citar quer os desdobramentos das atividades serão apresentados a seguir conjuntamente à coleta e resultados obtidos⁴.

Fase	Objetivo da tarefa	Breve descrição da Tarefa
(1) Levantamento	Escopo da Ontologia	Conhecer a audiência; Conhecer a ontologia de alto nível que fundamenta a ontologia em desenvolvimento e; Obter informações básicas sobre o contexto.
	Conhecer os fundamentos	Conhecer o escopo básico do domínio e respectivos conceitos.
	Identificar <i>expertise</i>	Identificar a <i>expertise</i> dos especialistas envolvidos
(2) Contato	Obter conhecimento	Elicitar conhecimento de especialistas
	Esclarecer dúvidas e mal entendidos	Esclarecer dúvidas ou mal entendidos
	Levantar e registrar termos	Registrar termos elicitados candidatos à ontologia
(3) Validação	Validar conhecimento	Obter aprovação sobre termos adquiridos e suas definições
	Atualizar	Atualizar dados após validação

Quadro 1 - Roteiro metodológico proposto

FONTE: Desenvolvido pelos autores

A atividade de AC, conforme roteiro proposto é apresentada a seguir como segue.

⁴ Detalhes acerca de cada uma das tarefas apresentadas se encontram em COELHO, Kátia C. Aquisição de conhecimento especializado para construção de ontologias: um estudo no domínio das ciências da vida. Dissertação de mestrado – Universidade Federal de Minas Gerais, Escola de Ciência da Informação, 2012.

4. COLETA DE DADOS E RESULTADOS

(1) LEVANTAMENTO: Na fase de Levantamento os dados obtidos das etapas: Escopo da ontologia; Escopo básico do domínio e Identificação da expertise dos especialistas se apresentam como segue:

Escopo da ontologia: Para se entender uma ontologia algumas sugestões são relevantes (ALMEIDA, 2006): i) conhecer a audiência; ii) conhecer o escopo planejado; iii) conhecer a ontologia de alto nível que fundamenta a ontologia; iv) obter informações básicas sobre a equipe e o contexto. O desenvolvimento de cada uma dessas etapas segue conforme apresentado a seguir:

Audiência: Para os propósitos da presente pesquisa, a ontologia em desenvolvimento – *Blood Ontology* (BLO) ⁵ se propõe a fornecer subsídios a: i) Vocabulários biomédicos: organização e publicação de vocabulário sobre hematologia e hemoterapia para fins de pesquisa e treinamento; ii) Sistemas de Informação: criação de condições que sirvam de referência para modelos de sistemas de informação; iii) troca de conhecimento entre especialistas; iv) fonte de conhecimento de novos pesquisadores (ALMEIDA; PROIETTI; SMITH, 2011).

Ontologia de alto nível que fundamenta a ontologia em desenvolvimento: de acordo com o definido no *Blood Project* (ALMEIDA; PROIETTI; SMITH, 2011), a *Blood Ontology* - BLO se fundamenta nos princípios da *Basic Formal Ontology* (BFO), uma ontologia de alto-nível criada em 2002 para apoiar pesquisas científicas. Ontologias de alto nível descrevem conceitos gerais como espaço, tempo, matéria, objeto, evento, ação, para citar alguns, os quais não dependem do problema ou domínio (ARP; SMITH, 2008).

Informações básicas sobre o contexto: A BLO compreende um conjunto de ontologias, em que cada uma engloba um grupo de questões relevantes na área de hematologia e hemoterapia. A BLO foi concebida considerando as práticas para ontologias descritas em Smith et al (2004) e fundamentada nos princípios preconizados pela BFO.

Escopo básico do domínio: Procurou-se mapear o domínio em estudo, a partir da literatura científica produzida sobre o assunto, no qual se baseou principalmente nas publicações do especialistas do GIPH. Além disso, foi

⁵ Disponível na Internet em :<<http://mba.eci.ufmg.br/BLO/>>. Acesso em: 20/03/2011.

referência para levantamentos iniciais acerca dos principais conceitos e endosso para os especialistas entrevistados. Efetuou-se uma consulta no Portal de Periódicos Capes. Obteve-se 1.802 trabalhos produzidos no período de 1985 a 2011. Esse recorte temporal foi determinado pelo resultado apresentado na consulta. O resultado mostra o Japão com 568 trabalhos, seguido dos Estados Unidos, com 435 trabalhos e o Brasil em seguida como terceiro país em publicações, com 188 trabalhos. No contexto brasileiro, o GIPH é responsável por 79 trabalhos produzidos sobre o assunto, considerando-se artigos em periódicos e anais em congressos. O artigo “*Global epidemiology of HTLV-I infection and associated diseases*” (PROIETTI, F.A.; CARNEIRO-PROIETTI, A.B. F; CATALAN-SOARES, B.C; MURPHY, E.L, 2005.) foi citado 188 vezes desde a sua publicação e identificado com um dos dez artigos mais citados.

Após o levantamento, procedeu-se à identificação do domínio do presente trabalho. O assunto central abordado era acerca do vírus linfotrópico de células T humanas (HTLV I/II) e respectivas doenças associadas. Considerando que o conhecimento do domínio como um todo é inviável, a presente pesquisa se limitou ao domínio da doença associada ao HTLV I, denominada Mielopatia Associada ao HTLV/Paraparesia Espástica Tropical (HAM/TSP).

Pos sua vez, para os entendimentos sobre conceitos básicos do domínio, o material produzido pelo GIPH foi estudado, o que possibilitou familiaridade mínima com o assunto, de modo a tornarem possíveis entendimentos iniciais relativos aos conceitos básico, sem a finalidade de extração de termos, uma vez que tal procedimento se daria a partir dos especialistas. Para tal, foi feita uma imersão na literatura procedendo-se à leitura de artigos e material fornecido pela coordenadora do GIPH.

Expertise dos especialistas do GIPH: A terceira tarefa da fase de levantamentos consistiu em identificar a *expertise dos especialistas* envolvidos. Levou-se em conta as observações de Swales (1987) para os requisitos relacionados à identificação de comunidades de pesquisa a seguir: i) *metas comuns*: refletidos nos documentos científicos produzidos pelo GIPH, cuja experiência encontra-se registrada na sua produção científica desde 1985; ii) *mecanismos de participativos*: cujas formas de comunicação se baseiam nas

reuniões para encontros sistemáticos, para troca de experiências e relatos de casos, entre outros; iii) *troca de informação*: cujos mecanismos utilizados para prover troca de informação com propósitos definidos são feitos por meio de simpósios e seminários; iv) *terminologia especializada*: identificado pelo vocabulário específico sobre HTLV e respectivas doenças; v) *alto nível de especialização*: especialistas com pós-doutorado, doutorado e mestrado nos temas correlatos ao HTLV. O GIPH é o único grupo de Minas Gerais que desenvolve pesquisa sobre este assunto.

(2) CONTATO: A fase de *Contato* consistiu de encontros com especialistas do domínio para fins de eliciação de conhecimento. Esta etapa visa à eliciação do conhecimento, cujas tarefas são respectivamente: Entrevistas com especialistas do domínio, Esclarecimento de dúvidas e mal entendidos e levantamento e registro de termos, conforme descrito a seguir.

Entrevista: O roteiro da entrevista era apresentado ao especialista. Explicações iniciais eram repassadas de modo a contextualizá-lo no objetivo da atividade. Nessa fase, a entrevista que era gravada visava obter termos relativos ao domínio da doença HAM/TSP.

Por sua vez, na tarefa relacionada à **esclarecer dúvidas e mal entendidos**, empregou-se a técnica de *teachback*. Antes disso, após cada sessão de eliciação era realizada uma análise do registro conforme exposto por cada um dos especialistas. Buscava-se inspecionar o registro gerado, no intuito de verificar alguma inconsistência com relação ao que ficou registrado. Finda a análise, o especialista era novamente contatado para novo encontro, para que os registros gerados e analisados pudessem ser checados pelo especialista. Era apresentado novamente ao especialista o que se encontrava registrado no *Protegé-frames*, para esclarecimentos acerca de termos e expressões, que eram reescritos ou acrescentados conforme o especialista desejava.

Na tarefa de **levantar e registrar termos**, os registros gerados pelos especialistas foram então analisados e selecionados. Esta análise visou identificar termos significativos ao domínio que por sua vez, seriam candidatos à ontologia. Esses termos foram extraídos e registrados preliminarmente numa

estrutura criada de modo que pudessem ser validados em seguida pelos especialistas.

(3) VALIDAÇÃO: A terceira etapa de AC visou validar os termos candidatos à ontologia e suas respectivas definições. Após a organização preliminar dos termos, procedeu-se à sua apresentação ao coordenador do GIPH para que pudesse validá-la. Nesta etapa, o especialista tinha o papel de aceitar ou sugerir mudanças no que havia sido registrado. Objetivava-se verificar principalmente se tais termos representariam o domínio da HAM, de acordo com o que foi apresentado pelo especialista do domínio durante a entrevista. A estrutura apresentada no Quadro 2 representa o conjunto de termos organizados após a validação do especialista.

Processo Etiológico	Infecção pelo HTLV-1
produz	
Desordem	Proliferação do provírus e de partículas virais nas Células T, com produção de substâncias virais do tipo TAX, REX, ENV e outras proteínas.
leva	
Disposição (doença)	Mielopatia Associada ao HTLV/Paraparesia Espástica Tropical (HAM/TSP)
realizada no	
Processo Patológico	Ação da proteína TAX nos linfócitos TCD4 e TCD8 positivas, nas células dendríticas e nas células cerebrais produzindo processo inflamatório crônico que afeta o sistema nervoso central levando a uma degeneração irreversível das células do sistema nervoso central.
produz	
Características corporais anormais	Atrofia da medula torácica com espessamento leptomeníngeo e atrofia medular
reconhecidas como	
Sintomas	<p>Espasticidade; fraqueza dos membros inferiores; distúrbio vesical; constipação intestinal; impotência; diminuição da libido; sintomas sensitivos (formigamento, agulhadas e queimação); dor lombar baixa com irradiação dos membros inferiores.</p> <p>Dificuldade para urinar; alterações intestinais; dificuldade para subir e descer escadas; alteração dos reflexos; alteração das sensações térmicas; alterações oftalmológicas; depressão; dificuldade para caminhar; dormência.</p>
Sinais	Sensibilidade vibratória diminuída; hiperreflexia dos membros inferiores; hiperreflexia dos membros superiores; sinais de Hoffmann e Trömer positivos; reflexo mandibular exaltado; clônus; sinal de babinsk positivo. Escamação da pele; alteração dos reflexos. Aumento dos reflexos; incontinência urinária; alteração da marcha; marcha espástica; hiperreflexia (aumento dos reflexos medulares); alteração da marcha. hiperreflexia ósteotendínea (patelar) e aquiliana; alterações da sensibilidade ao tato; alteração da sensibilidade à dor; fraqueza em diversos grupos musculares

	dos membros inferiores.
--	-------------------------

usados em

Processos interpretativos

produz

Hipóteses	<p>Descartar doenças que se assemelham à HAM/TSP mas que não seja HAM. Descartar doenças neurológicas que não sejam HAM</p> <p>Descartar mielopatias associadas idênticas à HAM/TSP, mas que não são positivas para o HTLV.</p> <p>Descartar outras doenças inflamatórias sobre a medula;</p> <p>Descartar doenças degenerativas da coluna vertebral;</p> <p>Descartar doenças tumorais;</p> <p>Descartar doenças parasitárias (esquistossomose);</p> <p>Descartar doenças vasculares.</p>
------------------	--

sugerem

Testes laboratoriais	<p>Testes sorológicos (soro): ELISA (teste de triagem); Western Blot (teste confirmatório); Teste no líquido: dosagem do líquido (líquido espinhal); Teste de PCR; Ressonância da medula espinhal; Tomografia; Exame radiológico para achar uma atrofia do cordão medular; Ultrassom do trato urinário; Exames de urodinâmica.</p>
-----------------------------	--

produzem

Resultados de testes	<p>Sorologia positiva para ELISA; Positiva para Western Blot e no Líquor que mostra um aumento discreto de célula, aumento discreto de proteína com presença ou não de banda; Presença de anticorpos anti-HTLV -1 no soro e no líquido (LCR); detecção do DNA proviral no sangue e/ou no líquido;</p> <p>Resultado da ressonância apontando uma lesão na medula espinhal;</p> <p>Ressonância do cérebro mostra uma alteração de sinal na substância branca da parte anterior do cérebro, do giro frontal;</p> <p>imunoeletroforese de proteínas, uma imunofixação que avalia os picos de proteínas IGG, IGM; PCR reação em cadeia de polimerase em tempo real detecta partícula viral no sangue periférico; Imagem mostra lesão na coluna; Presença de <i>flower cell</i> no líquido; Imagem: degeneração da medula.</p>
-----------------------------	--

usados em

Processos interpretativos

produzem

Resultados	<p>Diagnóstico que paciente X tem uma doença neurológica caracterizada como uma mielopatia e sorologia positiva para HTLV, conhecida como Mielopatia Associada ao HTLV/Paraparesia Espástica Tropical/TSP.</p>
-------------------	--

Quadro 2 - Validação pelos especialistas

FONTE: Desenvolvido pelos autores

Com relação ao resultado final apresentado ao especialista, verificou-se que foram propostas poucas mudanças nos termos considerados candidatos à ontologia. Além disso, numa análise preliminar da aplicação do roteiro metodológico - mesmo que necessite ser testado e melhorado em outros estudos, verificou-se a sua viabilidade para elicitación de conhecimento em domínios correlatos às ciências da vida. Tal constatação se deve ao fato de ter sido possível estabelecer diálogo mínimo com os especialistas para se levantar os principais termos relativos à doença HAM com o qual se estabeleceu o domínio. Essa estrutura se mostrou necessária e importante para conduzir o processo, inclusive por levar em consideração o tempo disponível para a realização tarefa com o especialista.

Entretanto, o presente trabalho apresenta algumas constatações acerca de dificuldades percebidas ao longo do processo da atividade de AC, a saber: i) tempo: considerado a principal dificuldade para participação de uma atividade de AC. De modo geral, o especialista tem pouco tempo disponível por acumular outras atividades como: ensino, coordenação, atendimento clínico, para citar alguns; ii) vocabulário: o vocabulário especializado dificulta a compreensão das terminologias e repassar conceitos complexos para quem não é da área. Tal fato foi evidenciado também por todos os entrevistados; iii) especialização: O assunto é multidisciplinar e os especialistas atuam em subdomínios: existem poucos especialistas para consulta e eles são super-especializados. Super-especialização no âmbito da própria especialidade, o que resulta em narrativas com diferentes níveis de granularidade para organizar a informação; iv) granularidade: definir o nível de granularidade para direcionar o processo de AC conforme o conhecimento de cada especialista.

5 CONSIDERAÇÕES FINAIS

Como resultado da presente pesquisa, espera-se que as proposições aqui apresentadas contribuam ao enriquecer a discussão sobre atividade de AC, principalmente em domínios especializados, além de abrir possibilidades de tal

discussão pela ciência da informação. Cabe destacar os desdobramentos técnicos e sociais dos resultados que buscam melhorias na forma de se registrar o conhecimento especializado. Espera-se que as discussões apresentadas contribuam, enriquecendo a discussão sobre a atividade de AC em domínios especializados, tornando seu trabalho mais efetivo e, no longo prazo, auxiliando o próprio especialista do domínio, que, no caso do presente trabalho, são médicos, biólogos, bioquímicos, para citar alguns.

Algumas questões ainda permanecem em aberto, nos quais cabe citar: o uso de uma ferramenta colaborativa, como a wiki, possibilita a efetiva participação de todos especialistas no domínio? O uso de uma wiki pode ser considerado um facilitador para as anotações diretamente na ontologia? O especialista deveria fazer anotações sozinho ou em conjunto com o cientista da informação? Além disso, identificar como as validações acontecem na prática e se existe alguma dificuldade por parte do especialista em fazer anotações, ou do cientista da informação em anotar o que de fato é obtido do especialista – o conhecimento registrado pelo engenheiro do conhecimento é o conhecimento do especialista?

A tarefa de aquisição de conhecimento, conforme apresentado no presente trabalho, é apenas parte do processo de construção de ontologias de determinado domínio. Como parte, é também responsável pela tarefa de apreensão, organização, formalização e compartilhamento do conhecimento, via ontologias, entre especialistas de um domínio do conhecimento.

AGRADECIMENTOS

Este trabalho foi desenvolvido com o apoio da Fundação Hemominas e pela Fundação de Amparo à Pesquisa de MG - Fapemig.

REFERÊNCIAS

ADAMS, F. Knowledge. In FLORIDI, L. (Ed.). *Philosophy of Computing and Information*. Malden: Blackwell. 2004. pp.228-237.

ALMEIDA, M. B. *Um modelo baseado em ontologias para representação da memória organizacional*. 2006. 321f. Tese (Doutorado em Ciência da Informação) - Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte.

BATES, M.J. The Invisible Substrate of Information Science. *Journal of the American Society for Information Science*. 50(12):1043-1050, 1999. Disponível em: <<http://mba.eci.ufmg.br/downloads/inforscience.pdf>>. Acesso: Mar. 2011.

BOOSE, J. H. Knowledge acquisition tools, methods, and mediating representations. In: *Japanese Knowledge Acquisition for Knowledge-based systems*. Workshop, JKAW, 1. 1990, Kyoto, Japan. Disponível em: <<http://pages.cpsc.ucalgary.ca/~gaines/BooseBradshaw/>>. Acesso: Abr. 2010.

BOOSE, J. H; GAINES, B. R. Knowledge Acquisition for Knowledge-Based Systems: Notes on the State-of-the-Art. 1989. *Machine Learning*, v. 4, p. 377-394. Disponível em: <<http://www.springerlink.com/content/n40354358u6j7903/>>. Acesso em: Jun 2010.

BURGE, J. E. *Knowledge Elicitation for design task sequencing knowledge*. 1998. 154 f. Degree of Master of Science. Worcester Polytechnic Institute. US. Disponível em: <<http://www.wpi.edu/Pubs/ETD/Available/etd-101399-123113/unrestricted/jburge.pdf>>. Acesso: Maio 2010.

CAMPBELL K. E. *et al*. Representing thoughts, words, and things in the UMLS. *Journal of the American Medical Information Association*, v. 5, n. 5, p. 421-31, 1998. Disponível em: <<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC61323/?tool=pubmed>> . Acesso: Ago. 2010.

CHOO, C. W. *A Organização do Conhecimento*. São Paulo: Ed. SENAC, 2003. 421 p.

COMPTON, P.; JANSEN, R. A philosophical basis for knowledge acquisition. Knowledge Acquisition. 3rd *European knowledge acquisition for knowledge based systems workshop*. 1989. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.31.7340>>. Acesso: Jul. 2010.

GANDON, F. *Distributed artificial intelligence and knowledge management: ontologies and multi-agent systems for a corporate semantic web*. 2002. 483f. Tese (Scientific Philosopher Doctorate Thesis in Informatics) – Doctoral School of

Sciences and Technologies of Information and Communication. INRIA and University of Nice, Nice, 2002.

FERNANDEZ, M; GOMEZ-PEREZ, A; JURISTO, H. *Methontology: From Ontological Art Towards Ontological Engineering*, 1997. Disponível em:

<<http://www.aaai.org/Papers/Symposia/Spring/1997/SS-97-06/SS97-06-005.pdf>>. Acesso: Ago. 2010.

FONSECA, F. The Double Role of Ontologies in Information Science Research. *Journal of the American Society for Information Science and Technology*, v.58, n. 6, pp. 786-793. 2007.

HARRIS, Z. On a theory of Language. *The Journal of Philosophy*, v. 73, n. 10, p. 253-276 1976. Disponível em <<http://www.jstor.org/stable/2025530>>. Acesso: Out. 2010.

HAWKINS, D. An analysis of expert thinking. *International Journal of Man-Machine Studies*. v. 18, p. 1-47, Jan. 1983. Disponível em:

<<http://www.sciencedirect.com/science/journal/00207373>>. Acesso: Ago. 2010.

HAYES-ROTH, F; WATERMAN D, A; LENAT, D.B. *Building Expert Systems*. Massachusetts: Ed. Addison-Wesley. 1983, 350 p.

HOEHNDORF, R *et al.* BOWiki: an ontology-based wiki for annotation of data and integration of knowledge in biology. *BMC Bioinformatics* 2009, 10(Suppl 5):S5. Disponível em: <<http://www.biomedcentral.com/1471-2105/10/S5/S5>>. Acesso: Set. 2010.

HOFFMAN, R, R. *et al.* Eliciting knowledge from experts: A methodological analysis. *Organizational Behavior and Decision Processes*. V. 62, n.2, May 1995. pg 129-158. Disponível em: <<http://hfs.sagepub.com/content/50/3/481.refs>>. Acesso: Jun. 2010.

HUA, J. Study on Knowledge Acquisition Techniques. *Second International Symposium on Intelligent Information Technology Application*. 2008. Disponível em:

<<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4739560>>. Acesso: Maio 2010.

KELLY, G. *Princípios da Psicologia dos Construtos Pessoais*. New York: Norton, 1955. New York: Norton, 1955.

LANCASTER, F. W. *Construção e uso de tesouros: curso condensado*. Brasília: IBICT, 1987. 106 p.

MASTELLA, L. S. *Um modelo de conhecimento baseado em eventos para Aquisição e Representação de Seqüências Temporais*. 2005. Dissertação (Mestrado em Ciência da Computação) – Instituto de Informática, UFRS, Porto Alegre.

MILTON, N.; CLARKE, D.; SHADBOLT, N. Knowledge engineering and psychology: Towards a closer relationship. *International Journal of Human-Computer Studies*, v. 64, n. 12, p. 1214-1229. 2006.

MILTON, N. R. *Knowledge acquisition in practice: a step-by-step guide*. Cranfield: Springer, 2007. 176p.

NATIONAL I. S.O. *Guidelines for the construction, format, and management of monolingual controlled vocabularies*, 2005. (ANSI/NISO Z39.19-2005). Disponível em: <<http://www.niso.org>>. Acesso: Abr. 2011.

NENADIC, G *et al.* Terminology-driven literature mining and knowledge acquisition in biomedicine. *International Journal of Medical Informatics*, v.67, n.1 p. 33-48, 2002.

NEWELL, A; SIMON, H.A. *Computer science as empirical inquiry: symbols and search*. 1975. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.104.2482&rep=rep1&type=pdf>>. Acesso: 2 ago. 2010.

NILSSON, N. *The Physical Symbol System Hypothesis: Status and Prospects*. Disponível em: <<http://ai.stanford.edu/~nilsson/OnlinePubs-Nils/PublishedPapers/pssh.pdf>>. Acesso: 20 set. 2010.

NONAKA, I. A Dynamic Theory of Organizational Knowledge Creation, Tokyo, *Organization Science*, v. 5, N. 1, Feb. 1994. Disponível em: <<http://citeseer.ist.psu.edu/viewdoc/download?jsessionid=57E73EC388700A1FF87FCCB5FA9E3C21?doi=10.1.1.115.2590&rep=rep1&type=pdf>>. Acesso: Jun 2010.

OGDEN, C. K. *O significado de significado: um estudo da influencia da linguagem sobre o pensamento e sobre a ciência do simbolismo*. Rio de Janeiro: Zahar, 1972. 349p.

PAYNE, P,R et al. Conceptual knowledge acquisition in biomedicine: a methodological review. *J Biomedical Information*, 2007. v 40, n. 5, p. 82–602. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/17482521>>. Acesso: Ago 2010.

SCHEUERMANN, R et al. Toward an ontological treatment of disease and diagnosis. *Proceeding of the 2009 Summit on Translational Bioinformatics*, San Francisco, CA, pp. 116-120.

SHADBOLT, N. et al. Laddering: technique and tool use in knowledge acquisition. *Knowledge Acquisition*, p. 315-341, 1994. Disponível em: <<https://blog.itu.dk/SLR-F2010/files/2010/07/paper-1-pages-1-12-15.pdf>>. Acesso: Ago 2010.

SHADBOLT, N. Eliciting Expertise. *Evaluation of Human Work*. Ed. Taylor & Francis. 2005. Disponível em: <http://eprints.ecs.soton.ac.uk/14563/1/Eliciting_Expertise.pdf>. Acesso: Jul 2010.

SHADBOLT, N.; SWALLOW, S. *Epistemics: Knowledge Acquisition*. (online) Disponível em: <<http://www.epistemics.co.uk/Notes/63-0-0.htm>>. Acesso: Jul 2010.

SMITH, B. New Desiderata for Biomedical Terminologies. In Munn, K.; Smith, B. (Ed.). *Applied Ontology*. Frankfurt: Ontos Verlag. 2008. pp. 21 -39.

STEHR, H et al. PDBWiki: added value through community annotation of the Protein Data Bank. *Oxford Journals Life Sciences*, 2010. Disponível em: <<http://database.oxfordjournals.org/content/2010/baq009.full>>. Acesso: Set. 2010.

SOUZA, R.R; ALMEIDA, M.B. Representação do conhecimento: identidade ou esvaziamento da Ciência da informação? *IV Encontro Ibérico EDIBCIC 2009*, Coimbra, Portugal. Disponível em: <<http://mba.eci.ufmg.br/downloads/EDIBCIC2009CoimbraFinal3a.pdf>>. Acesso: Abr. 2011.

TSUMOTO, S. Automated Knowledge Acquisition from Clinical Databases based on Rough Sets and Attribute-Oriented Generalization. *American Medical Informatics Association Symposium*. p. 548–552, 1998. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.138.1420>>. Acesso: Ago. 2010.

TURBAN, E. *Expert systems and applied artificial intelligence*. New York: Macmillan Publishing Company, 1992. 804 p.

VAN-AUKEN, K et al. Semi-automated curation of protein subcellular localization: a text mining-based approach to Gene Ontology (GO) Cellular Component curation. *BMC Bioinformatics*, 2000. Disponível em: <<http://www.biomedcentral.com/1471-2105/10/228>>. Acesso: Set. 2010.

VICKERY, B. C. Knowledge representation: a brief review. *Journal of Documentation*, London, v. 42, n. 3, p.145- 159, Sept. 1986.

VITA, R et al. Curation of complex, context-dependent immunological data. *BMC Bioinformatics*, v. 7, 2006. Disponível em: <<http://www.biomedcentral.com/1471-2105/7/341>>. Acesso: Set. 2010.

WANG, Y et al. Knowledge elicitation plug-in for Protégé: Card sorting and laddering. In: *The Semantic Web, ASWC*, v. 4185, p. 552-565, 2006. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.71.7546&rep=rep1&type=pdf>>. Acesso: Set. 2010.

WOLF, R.; DELUGACH, H. S. *Knowledge Acquisition via tracked repertory grids*. Computer Science Dept. Univ. Alabama in Huntsville, 1996. (relatório técnico). Disponível em: <<http://www.cs.uah.edu/tech-reports/TR-UAH-CS-1996-02.pdf>>. Acesso: Jun. 2010.

