



Um modelo para a visualização do conhecimento baseado em arquétipos visuais

Héctor Andrés Melgar Sasieta^{1*}, Fabiano Duarte Beppler² and Roberto Carlos dos Santos Pacheco³

¹Departamento de Ingeniería, Sección de Ingeniería Informática, Pontificia Universidad Católica del Perú, Av. Universitaria, 1801, San Miguel, Lima 32, Peru. ²Instituto Stela, Florianópolis, Santa Catarina, Brazil. ³Programa de Pós-graduação em Engenharia e Gestão do Conhecimento, Universidade Federal de Santa Catarina, Florianópolis, Santa Catarina, Brazil. *Autor para correspondência. E-mail: amelgar@pucp.edu.pe

RESUMO. Neste artigo apresenta-se um modelo que visa facilitar a visualização do conhecimento armazenado em repositórios digitais usando arquétipos visuais. Os arquétipos são estruturas que contêm representações visuais do mundo real que *a priori* são conhecidas pelo grupo-alvo, e possuem estruturas semânticas que permitem identificar os conceitos do domínio representados em cada região. O modelo proposto apoia-se no *framework* para visualização do conhecimento proposto por Burkhard e descreve as interações dos usuários com os arquétipos visuais. O usuário por meio dos arquétipos pode recuperar e visualizar o conhecimento relacionado aos conceitos representados nas imagens dos arquétipos. Um protótipo foi desenvolvido para demonstrar a viabilidade do modelo usando arquétipos no domínio da anatomia, a *Foundational Model of Anatomy* e a *Unified Medical Language System* como conhecimento do domínio e o banco de dados da *Scientific Electronic Library Online* como repositório de documento. O uso das representações visuais nos arquétipos facilita a divulgação do conhecimento já que ao fazer parte da visão do mundo dos usuários, podem facilmente ser associadas a conhecimentos prévios. As representações visuais são processadas rapidamente no cérebro requerendo menos esforço que o processamento de informação textual.

Palavras-chave: visualização do conhecimento, recuperação do conhecimento, anotação semântica, ontologia.

A model for knowledge visualization based on visual archetypes

ABSTRACT. This paper presents a model that aims to facilitate the visualization of the knowledge stored in digital repositories using visual archetypes. Archetypes are structures that contain visual representations of the real world that are known *a priori* by the target group, and which have semantic structures for identifying the entities of the domain represented in each region. The proposed model is supported by the framework for knowledge visualization proposed by Burkhard and describes the users' interactions with visual archetypes. The user through the archetypes can retrieve and view the knowledge related to the entities represented in the archetypes' images. A prototype was developed to demonstrate the feasibility of the model using archetypes in the biomedical field, the Foundational Model of Anatomy and the Unified Medical Language System as domain knowledge and the Scientific Electronic Library Online database as a document repository. The use of visual representations in archetypes facilitates the dissemination of knowledge, because these are part of the world view of users and can easily be related with prior knowledge. Visual representations are processed quickly in the brain and require less effort than the processing of textual information.

Keywords: knowledge visualization, knowledge retrieval, semantic annotation, ontology.

Introdução

A divulgação e o compartilhamento de conhecimento são tarefas complexas para organizações uma vez que elas, muitas vezes não sabem o que têm e não possuem sistemas que permitam localizar e recuperar eficientemente o conhecimento que nelas reside (ALAVI; LEIDNER, 2001). Uma quantidade considerável de conhecimento explícito encontra-se espalhado em vários documentos dentro das organizações. Em muitos casos, a possibilidade de acessar de forma

eficiente (i.e., recuperar) e reutilizar este conhecimento é limitada.

Sistemas de visualização de informação podem ser usados para explorar o conhecimento existente dentro de repositórios de informação, navegando sobre grandes volumes, examinando os dados para fazer novas descobertas ou ganhar *insights* (GERSHON et al., 1998). Este tipo de sistema é especialmente útil quando as pessoas necessitam de alguma informação, mas elas não conseguem traduzir essas necessidades em termos de busca a serem postados em sistemas de recuperação de

informação (RI) tradicionais. Os sistemas de visualização de informação ajudam a restringir mais rapidamente o domínio da busca o que torna o processo de encontrar dados específicos mais fácil e dinâmico (FEKETE et al., 2008).

Nesse contexto, o campo da visualização do conhecimento tem pesquisado como o uso de elementos visuais pode ajudar no processo de transferência do conhecimento. O uso de imagens que contêm representações do mundo real, ou seja, que fazem parte da visão do mundo dos usuários permite que o conhecimento apresentado por meio destas imagens possa facilmente ser relacionado com conhecimentos prévios, facilitando assim a disseminação do conhecimento. Neste trabalho, este tipo de imagem é encapsulado em uma estrutura denominada arquétipo visual.

Esta pesquisa apresenta um modelo que visa facilitar a visualização do conhecimento usando arquétipos visuais. O objetivo do modelo proposto consiste na utilização de arquétipos visuais, que não apenas contenham uma representação visual do mundo, mas também anotações semânticas que ajudam a descrever essa representação visual, como estrutura de apoio ao processo de visualização do conhecimento. A ideia é que o usuário ao visualizar um arquétipo possa rapidamente reconhecer quais regiões possuem conhecimento associado e possa recuperar os documentos relacionados às regiões selecionando apenas a região de interesse. Por exemplo, dado um arquétipo que represente o motor de um carro formado por uma imagem contendo cilindros, pistões e válvulas, para recuperar os documentos relacionados aos pistões, bastará selecionar a região que representa os pistões.

Este artigo está estruturado da seguinte maneira: após esta introdução, é apresentada a revisão da literatura sobre os tópicos de visualização do conhecimento e anotação semântica de imagens. Posteriormente, é descrito o modelo proposto para a visualização do conhecimento. Nas seções seguintes, são apresentados o material e os métodos usados para o desenvolvimento do modelo e a discussão. Finalmente, na última seção, apresentam-se as conclusões.

Revisão da literatura

Visualização do conhecimento

A visualização do conhecimento pode ser definida como o uso de representações visuais para melhorar a transferência de conhecimentos entre pelo menos duas pessoas ou grupo de pessoas (BURKHARD, 2005; EPPLER; BURKHARD, 2007). Ora, tornar visível o conhecimento para que

ele possa ser acessado, discutido, avaliado, apreciado ou gerenciado, é um objetivo de longa data na gestão do conhecimento. Por tais aspectos a visualização do conhecimento tornou-se recentemente o foco de atenção nas comunidades acadêmica e de negócios (HOU; PAI, 2009). A visualização do conhecimento analisa métodos que enfrentam os desafios existentes dentro das organizações referentes à transferência do conhecimento, comunicação para diferentes tipos de usuários e sobrecarga de informação (BURKHARD; MEIER, 2005).

Os benefícios proporcionados pela visualização parece ser que dependem do fato de ela atuar como um frame de referência ou área de armazenamento temporária para os processos da cognição humana. A visualização potencializa a memória ao fornecer um extenso conjunto de trabalho para análise e reflexão, tornando-se assim um facilitador externo da cognição (FEKETE et al., 2008). Segundo Ware (2000), há, principalmente, duas teorias na psicologia que explicam como a visão pode ser usada efetivamente para perceber elementos e formas. No baixo nível, a teoria do processamento pré-atencional explica que alguns elementos visuais podem ser processados rapidamente. No nível mais alto, a teoria da forma ou teoria da Gestalt descreve alguns princípios usados pelo nosso cérebro para entender uma imagem.

O processo para visualizar informação pode ser dividido em duas etapas: na primeira delas, a informação é processada pelo olho e pelo córtex visual primário, onde neurônios individuais em áreas específicas (denominadas V1, V2, V3, V4, MT) são especializados para identificar características particulares (por exemplo: orientação, cor, textura, contorno ou movimento). Nesta fase a informação é processada de forma pré-atencional e muito rapidamente. Na segunda fase, o processamento da informação é dividido dentro de dois subsistemas complementares e independentes, um deles focado na identificação dos objetos (quê) e o outro focado na localização espacial (onde) (BURKHARD, 2005).

Os sistemas visualização de conhecimento são concebidos para que façam uso das habilidades que os humanos possuem para o processamento de imagens. As imagens são processadas pelo sistema nervoso antes do que o texto. Por outro lado, o uso de imagens que contêm representações do mundo real, que “a priori” são conhecidas pelo grupo-alvo e fazem parte da sua visão do mundo, permitem que o conhecimento apresentado por meio destas imagens possa facilmente ser relacionado com os conhecimentos prévios dos indivíduos, facilitando a aprendizagem e a recuperação (BURKHARD, 2005).

Visando orientar a aplicação da visualização do conhecimento nas organizações, Burkhard (2005) e Eppler e Burkhard (2007) propõem um *framework* baseado em cinco perspectivas que respondem a cinco perguntas-chave: 1) Que tipo de conhecimento precisa ser visualizado? 2) Por que o conhecimento deve ser visualizado? 3) A quem está sendo destinado? 4) Em que contexto deveria ser visualizado? e 5) Como pode o conhecimento ser visualizado? A listagem de possíveis respostas a estas perguntas fundamentais leva a um quadro conceitual que visa fornecer uma visão geral do campo de visualização do conhecimento e orientar a sua aplicação.

Anotação semântica de imagens

A anotação pode ser definida como o processo de tornar explícita a interpretação de um documento. A criação de metadados é uma das principais técnicas usadas para anotar documentos. Os metadados podem ser atribuídos a uma ampla variedade de documentos, pode ser expressa em diversas linguagens e vocabulários (CORCHO, 2006), e pode ser feita de forma manual, automática ou semiautomática (UREN et al., 2006).

Ontologias têm sido usadas para anotar documentos (CORCHO, 2006). Estruturas ontológicas acrescentam valor às anotações semânticas permitindo a realização de inferências e a navegação conceptual. Quando as ontologias são usadas, os documentos são etiquetados com descrições semânticas (*i.e.*, informações sobre classes e instâncias) (KIRYAKOV et al., 2004).

Os metadados associados às imagens podem ser classificados como: i) 'metadados com conteúdo' independente, quando estes se encontram relacionados à imagem, mas não a descrevem (*i.e.*, autores, data de criação etc.); ii) 'metadados com conteúdo dependente', quando se referem a características de baixo nível e/ou nível intermediário (*i.e.*, cor, textura etc.); iii) 'metadados com conteúdo descritivo', quando se referem a conteúdo semântico (*i.e.*, relações das entidades da imagem com entidades do mundo real) (HANBURY, 2008).

Os metadados com conteúdo descritivo podem ser especificados usando um ou mais dos seguintes enfoques: i) 'descrições usando texto livre', quando nenhuma estrutura pré-definida para a anotação é proporcionada; ii) 'classificação usando palavras-chave', quando palavras-chave selecionadas de forma arbitrária ou a partir de vocabulários controlados são usadas para descrever as imagens; iii) 'classificação baseada em ontologias', quando conceitos ou instância de uma ontologia são usados no processo de anotação (HANBURY, 2008; MÜLLER et al., 2004).

Por outro lado, os metadados com conteúdo descritivo podem ser proporcionados a dois níveis de

especificidade: i) conteúdo descritivo associado à imagem completa (SCHREIBER et al., 2001) e ii) segmentação da imagem com vinculação do conteúdo descritivo em cada região da imagem (HSU et al., 2009; RUBIN et al., 2009).

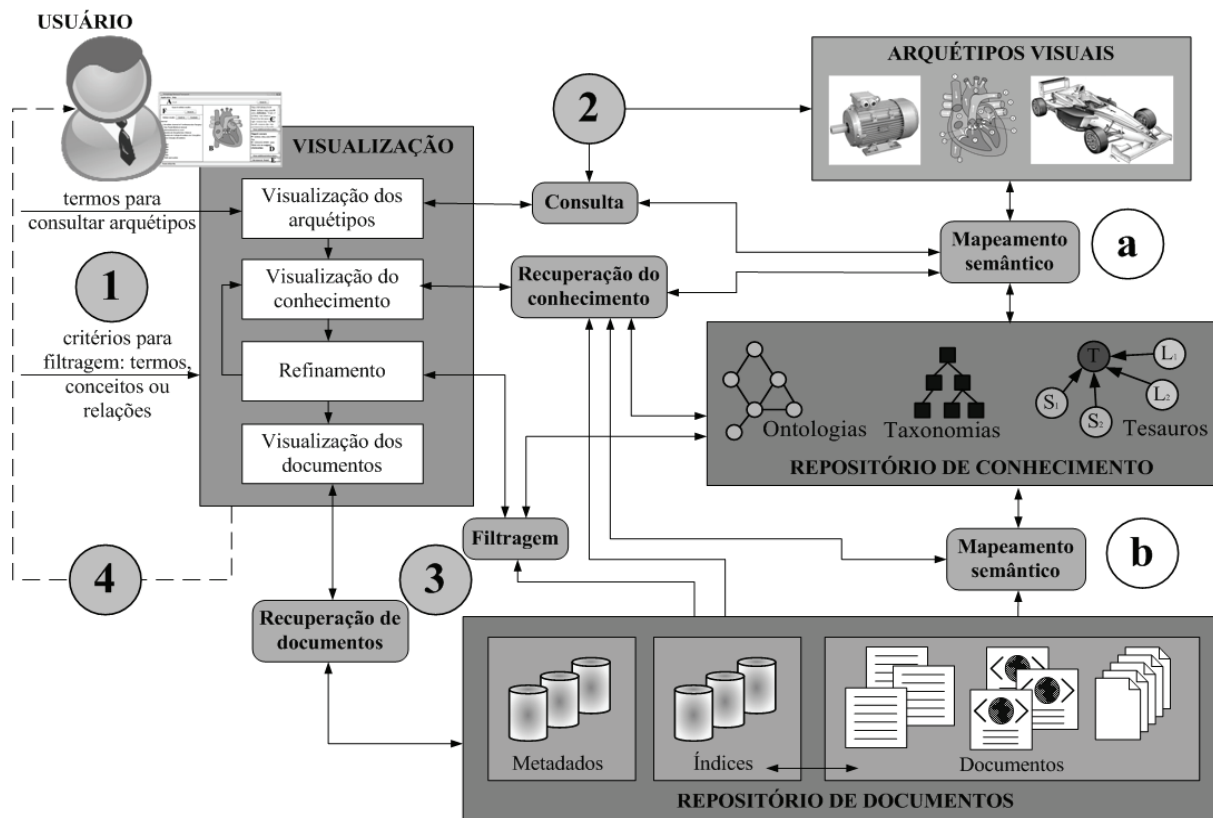
Os sistemas para anotar imagens usando metadados com conteúdo descritivo baseados em ontologias, geralmente usam dois tipos de ontologias, uma para definir o esquema de anotação e outra para definir os conceitos do domínio. Isto permite que os esquemas de anotação sejam definidos de forma independentes ao domínio do conhecimento. A forma de definir a ontologia de anotação depende dos requerimentos da aplicação. Por exemplo, Schreiber et al. (2001) usaram uma estrutura baseada em quatro elementos 'agente – ação – objeto – cenário' para anotar imagens de macacos.

Modelo para visualização do conhecimento

O modelo proposto foi concebido com o objetivo de facilitar a visualização do conhecimento usando como estruturas de apoio arquétipos visuais. O modelo conceitual, descrito na Figura 1 por quatro componentes: arquétipo visual, repositório de documentos, repositório de conhecimento e visualização.

Para que o conhecimento vinculado aos arquétipos seja recuperado de forma eficiente (*i.e.*, buscas semânticas), tanto os arquétipos quanto os repositórios de documentos devem ter sido previamente enriquecido com conteúdo semântico (letras a e b na Figura 1) obtido a partir do repositório de conhecimento (*i.e.*, ontologias, taxonomias, tesouros). Os mapeamentos semânticos dos arquétipos são realizados de forma manual tendo em consideração as necessidade e visão do mundo do grupo-alvo. Os mapeamentos semânticos dos repositórios são realizados de forma automática usando diversas abordagens como processamento simbólico, processamento da linguagem natural ou técnicas de linguística computacional.

O processo inicia quando o usuário precisa satisfazer alguma necessidade de informação. O usuário seleciona o arquétipo (número 1 na Figura 1) a partir do qual será visualizado o conhecimento. O critério de seleção dependerá diretamente das necessidades de informação do usuário, levando em consideração os conceitos representados nos arquétipos. Por exemplo, caso seja preciso visualizar o conhecimento associado ao coração, o usuário deverá selecionar um arquétipo no qual o coração encontre-se representado graficamente.



Fonte: elaborado pelos autores.

Figura 1. Modelo para a visualização do conhecimento.

Uma vez selecionado o arquétipo, o processo de visualização é executado (número 2 na Figura 1). O processo de visualização visa fornecer ao usuário o conhecimento armazenado nos repositórios de documentos.

A visualização dos resultados é realizada a partir dos nos arquétipos visuais, de forma que os usuários rapidamente percebam a quantidade de documentos associados a cada região do arquétipo. Esta representação é refletida no arquétipo, por exemplo, alterando a coloração das regiões nas quais se concentra a maior/menor quantidade de documentos ou incluir a quantidade de documentos recuperados sobre as regiões.

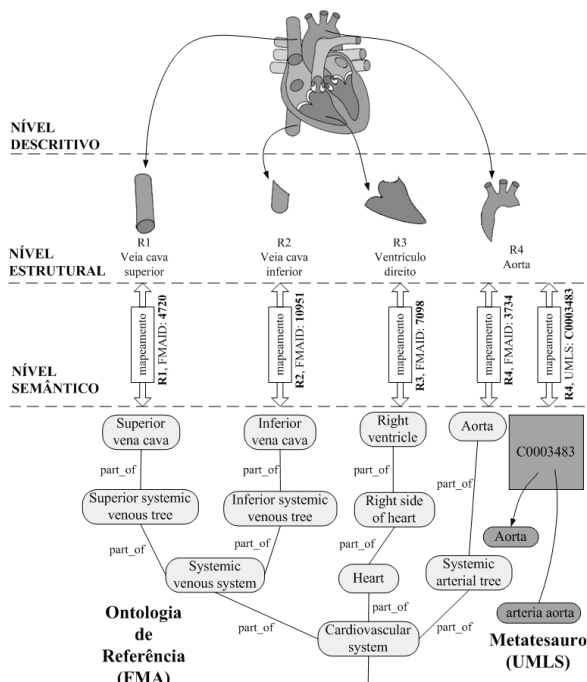
O usuário, a partir da visualização fornecida pelo arquétipo, pode restringir o espaço das buscas (número 3 na Figura 1) podendo usar para isto os metadados fornecidos pelos repositórios de documentos ou os conceitos definidos nos artefatos de conhecimento, iniciando assim um novo processo de visualização. A continuação se descreve brevemente cada um dos componentes.

Componentes do modelo

Arquétipo visual: um arquétipo visual pode ser definido como uma estrutura que permite

representar graficamente os conceitos de um domínio, no qual cada uma das partes que o compõem é especificada de forma explícita e formal, como por exemplo, em uma ontologia. É definido em três níveis, tal como pode ser apreciado na Figura 2: o ‘nível descritivo’ visa identificar e descrever as representações visuais; o ‘nível estrutural’ fornece informações sobre a estrutura interna visando explicitar as regiões dos arquétipos; e o ‘nível semântico’ tem como objetivo descrever os mapeamentos semânticos.

Repositório de documentos: este componente representa os repositórios a partir dos quais o conhecimento é recuperado. No modelo, os repositórios de documentos são caracterizados por possuírem uma camada semântica que permite formalizar a informação contida nos documentos (Figura 3) são definidos em quatro níveis: o ‘visa identificar’ o repositório; o ‘nível dos metadados’ visa descrever a estrutura da informação; o ‘nível do conteúdo’ fornece as estruturas necessárias para armazenar os documentos e os metadados. Semelhante aos arquétipos, o repositório de documentos também possui um ‘nível semântico’ que visar vincular conteúdo semântico aos documentos.

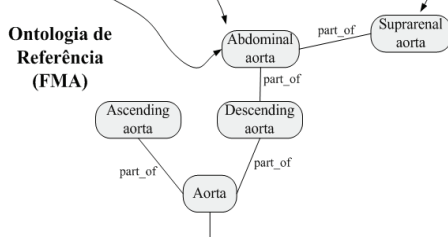


Fonte: elaborado pelos autores.

Figura 2. Composição dos arquétipos visuais.

BONAMIGO, Telmo Pedro and SIQUEIRA, Iara. *Screening for abdominal aortic aneurysms*. Rev. Hosp. Clin. [online]. 2003, v. 58, n. 2, pp. 63-68. ISSN 0041-8781.

OBJECTIVE AND METHODS: Screening for abdominal aortic aneurysms may be useful to decrease mortality related to rupture. We conducted a study to assess the prevalence of abdominal aortic aneurysms in southern Brazil and to define risk factors associated with high prevalence of this disorder. The screening was conducted using abdominal ultrasound. Three groups were studied: Group 1 - cardiology clinic patients; Group 2 - individuals with severe ischemic disease and previous coronary surgery, or important lesions on cardiac catheterism; Group 3 - individuals without cardiac disease selected from the general population. All individuals were male and older than 54 years of age. The ultrasonographic diagnosis of aneurysm was based on an anteroposterior **abdominal aorta** diameter of 3 cm, or on an **abdominal aorta** diameter 0.5 cm greater than that of the **supra-renal aorta**. **RESULTS:** A total of 2.281 people were screened for abdominal aortic aneurysms in all groups: Group 1 - 768 individuals, Group 2 ...



Fonte: elaborado pelos autores.

Figura 3. Anotação semântica dos documentos.

Repositório de conhecimento: este componente está conformado pelos artefatos usados para representar o conhecimento do domínio usado para realizar tanto o mapeamento semântico dos arquétipos e os documentos dos repositórios assim como para realizar as inferências, quando possíveis, sobre os conceitos usados nos processos de visualização do conhecimento.

Visualização: o componente de visualização é o encarregado de apresentar ao usuário os resultados das buscas. Este componente baseia-se nas seguintes tarefas: primeiro obter uma visão global dos dados (*overview first*), concentrar-se em itens de interesse e

filtrar itens irrelevantes, (*zoom and filter*) e finalmente fornecer detalhes sob demanda (*then details on demand*). A visão global dos dados é feita usando os arquétipos visuais – em que os elementos visuais são alterados visando que o usuário facilmente identifique as regiões onde se encontra conhecimento. A filtragem dos itens irrelevantes é facilitada pelos metadados dos repositórios e pelos mapeamentos semânticos. Para obter os detalhes sob demanda, o componente interage com os repositórios de documentos a fim de obter documentos específicos.

Material e métodos

Visando demonstrar a viabilidade de modelo proposto, foi desenvolvido um protótipo aplicado ao domínio biomédico. O repositório de conhecimento foi composto pela ontologia *Foundational Model of Anatomy* (FMA), uma ontologia de referencia no domínio anatômico (ROSSE; MEJINO, 2003, 2008), e pelo metatesauro *Unified Medical Language System* (UMLS), um repositório que integra vários vocabulários biomédicos (BODENREIDER, 2004).

Dado que a FMA foi desenvolvida usando *frames* de Protégé (<http://protege.stanford.edu/>) usou-se a API do Protégé para acessar as suas classes e propriedades. A UMLS, por outro lado, foi instalada em um banco de dados relacional. Como repositório de documentos usou-se uma cópia da base de dados SciELO (*Scientific Electronic Library Online*). SciELO é uma biblioteca eletrônica que abarca uma coleção seleta de revistas científicas brasileiras. Para anotar os artigos usou-se Metamap, um programa que identifica conceitos UMLS em textos escritos em linguagem natural. Metamap usa uma abordagem intensiva de conhecimento baseada no processamento simbólico, processamento da linguagem natural e técnicas de linguística computacional. Os arquétipos visuais foram implementados usando a tecnologia Java 2D. Para segmentar as regiões das imagens, usou-se AnnotImage, um *software* que permite anotar imagens usando a linguagem IML (*Image Markup Language*). AnnotImage, disponibiliza um XML por cada imagem segmentada, onde as regiões segmentadas são explicitamente especificadas. A anotação das imagens foi feita manualmente vinculando às regiões a conceitos da ontologia FMA e do metatesauro UMLS. Visando facilitar o processo de busca, todos os artigos foram indexados usando Lucene. Tanto o conteúdo textual (*i.e.*, título do artigo, resumo, palavras-chave etc.) quanto às anotações de 77.461 artigos foram indexadas.

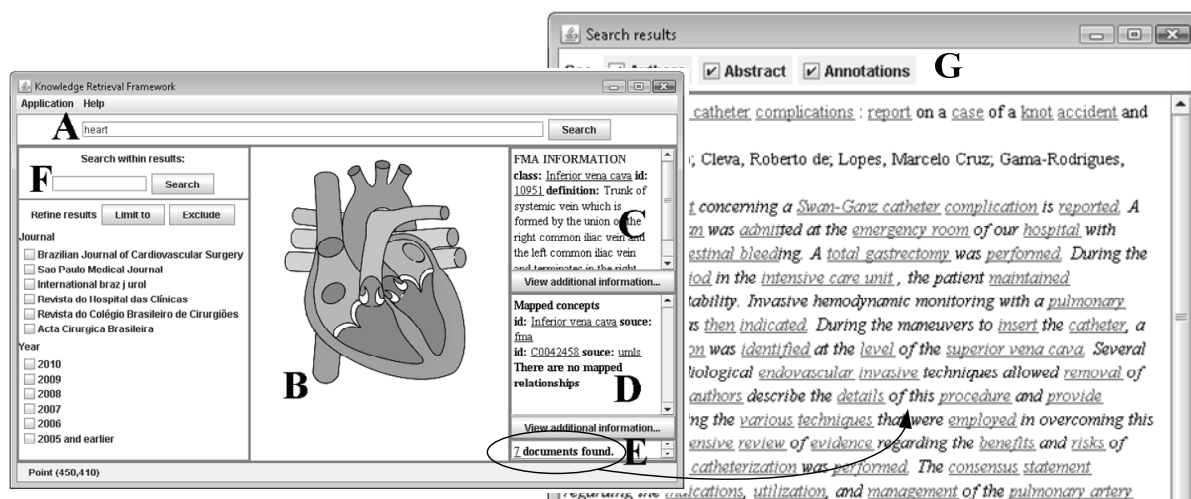
Na Figura 2 pode-se observar um exemplo de um arquétipo usado para visualizar conhecimento relacionado ao coração. Como elemento visual foi utilizado uma imagem em duas dimensões onde quatro regiões foram explicitadas: a veia cava superior, a veia cava inferior, o ventrículo direito e a aorta. Para anotar semanticamente as regiões, utilizou-se a FMA. Cada região foi vinculada a uma classe da FMA incorporando ao arquétipo todo o conhecimento explicitado nesta ontologia. Por exemplo, ao vincular a região R1 (etiquetada com veia cava superior) à classe 'Superior vena cava', o arquétipo agora 'entende' que aquela região é parte do sistema cardiovascular (classe *Cardiovascular system* na FMA) e que também é conhecida como 'Anterior vena cava' (sinônimo em inglês) e 'Vena cava superior' (nome equivalente em espanhol). A Figura 2 apresenta também como a região R4 (etiquetada como aorta) é mapeada ao conceito C0003483 da UMLS. Usando as relações UMLS, o arquétipo 'sabe' que *Aneurysm* é uma doença relacionada à aorta.

Na Figura 3 apresenta-se um exemplo de mapeamento do conteúdo de um documento usando também a ontologia FMA. Os benefícios que o mapeamento semântico oferece aos repositórios de documentos são os mesmos que os fornecidos aos arquétipos (*i.e.*, formalização do conhecimento, recuperação de informações relacionadas, execução de inferências etc.). Para o modelo, o fato dos arquétipos e os repositórios de documentos serem mapeados usando os mesmos artefatos permitem que estes dois componentes sejam integrados por meio das representações do conhecimento. Por exemplo, na Figura 2 a região R4 (etiquetada com aorta) encontra-se mapeada à

classe *Aorta* da FMA, por outro lado na Figura 3, observa-se como os conceitos 'Abdominal aorta' e 'suprarrenal aorta' foram identificados dentro do texto de um documento. Apesar de os mapeamentos não serem feitos usando as mesmas classes FMA, é possível inferir que o documento em questão possui conhecimento relacionado à região R4, já que a aorta suprarrenal (classe 'suprarrenal aorta' na FMA) é parte da aorta abdominal (classe 'Abdominal aorta' na FMA) que por sua vez é parte da aorta descendente (classe *Descending aorta* na FMA) que forma parte da 'Aorta'.

Na Figura 4 pode-se observar a interface principal do protótipo. Para procurar os arquétipos, o usuário ingressa os termos de busca na caixa de texto no painel superior (Figura 4-A). A ferramenta recupera todos os arquétipos associados aos termos processando a consulta tanto no nível estrutural (*i.e.*, busca textual pelo nome das regiões) quanto no nível semântico (*i.e.*, busca conceitual pelo mapeamento). Quando o processamento é feito no nível semântico, os termos são transformados em conceitos do domínio. Esta transformação permite que a consulta seja executada semanticamente, tornando a busca independente da linguagem. Por exemplo, para pesquisar arquétipos relacionados ao coração, o usuário pode inserir como termos de pesquisa palavras como coração, *corazón* (coração em espanhol), *heart* ou (coração em inglês). Em todos os casos, o processo retorna os mesmos resultados.

Depois de recuperar o arquétipo, este é colocado no painel central (Figura 4B). O usuário pode utilizá-lo de duas maneiras: para obter conhecimentos relacionados aos conceitos representados nas regiões ou para recuperar os documentos que mencionam os conceitos.



Fonte: elaborado pelos autores

Figura 4. Framework para a Recuperação do Conhecimento.

O conhecimento relacionado ao conceito é apresentado no painel direito no topo (Figura 4-C). Nesta versão do protótipo, esta informação é apresentada usando a FMA. Neste painel o usuário pode visualizar o nome, a identificação e a descrição da classe. Usando o botão *View additional information*, pode obter informações adicionais como nomes em outras línguas ou as entidades anatômicas que a constituem. Os conceitos mapeados em cada região podem ser observados no painel direito no centro (Figura 4-D). Esta informação é obtida a partir do nível semântico do arquétipo.

Ao selecionar uma região da imagem, além de apresentar ao usuário informações sobre o conceito, também é apresentada a quantidade de documentos relacionados com os conceitos mapeados em cada região. Esta informação está no painel direito, ao selecionar o vínculo (Figura 4-E), é apresentado um formulário contendo uma lista de todos os documentos do repositório relacionados a esse conceito. Na Figura 4-G pode-se visualizar uma lista dos documentos recuperados. Como cada documento está anotado, vínculos de cada conceito reconhecido nos documentos estão disponíveis, ao selecioná-los uma descrição do conceito é apresentada. O usuário pode refinar os resultados da pesquisa utilizando as opções disponíveis no painel esquerdo (Figura 4-F). Este painel apresenta duas opções, filtragem por revistas (i.e., *journals*) e filtragem por anos. O usuário pode procurar por texto dentro de documentos ou usar os filtros para refinar a busca.

Resultados e discussão

O resultado deste trabalho é um modelo que visa suportar a visualização do conhecimento usando arquétipos visuais como estruturas de apoio. Com o objetivo de validar o modelo proposto, foi desenvolvido um protótipo usando arquétipos na área da anatomia e repositórios de documentos no domínio biomédico.

No modelo proposto, tanto os arquétipos visuais quanto os repositórios de documentos têm sido enriquecidos com conteúdo semântico que possibilitam a integração destes dois componentes. O conteúdo semântico permite ao modelo 'entender' quais os conceitos representados tanto nos arquétipos quanto nos documentos para facilitar a busca conceitual. Uma das vantagens disto é a recuperação de conhecimento de forma independente à linguagem da escrita dos documentos. Por este 'entendimento', o modelo é capaz de recuperar os documentos relacionados aos conceitos representados no arquétipo auxiliando o

usuário no processo de busca. Este mecanismo de recuperação pode ser observado no protótipo e para recuperar documentos o usuário só precisa selecionar uma região do arquétipo sem necessidade de ingressar os termos das buscas como nos sistemas de recuperação tradicionais. Entretanto, o usuário pode ingressar termos para procurar arquétipos. Uma vez selecionado o arquétipo, o modelo 'entende' quais as informações necessitam ser recuperadas usando para tal os mapeamentos semânticos.

A ideia do modelo é que os resultados das buscas sejam apresentados de forma global no próprio arquétipo alterando a coloração das regiões onde se concentra a maior quantidade de documentos. Este comportamento do modelo é consistente com as tarefas de visualização (i.e., primeiro apresentar a visão global dos dados, depois concentrar-se em itens de interesse e filtrar itens irrelevantes e finalmente fornecer a informação sob demanda).

O modelo proposto apoia-se no *framework* para visualização do conhecimento proposto por (BURKHARD, 2005; EPPLER; BURKHARD, 2007). Neste *framework*, são propostas cinco perguntas-chave para orientar a aplicação da visualização do conhecimento. As respostas a estas perguntas, descritas na Tabela 1, forneceram o quadro conceitual que norteou o trabalho.

Tabela 1. Respostas das perguntas chaves do *framework* para visualização do conhecimento.

Pergunta-chave (<i>framework</i> para a visualização do conhecimento)	Resposta
Que tipo de conhecimento precisa ser visualizado?	O conhecimento armazenado nos repositórios de documentos.
Por que o conhecimento deve ser visualizado?	Para facilitar a divulgação do conhecimento.
A quem está sendo destinado?	A um grupo de pessoas que compartilham uma mesma visão do mundo.
Em que contexto deveria ser visualizado?	Em um entorno virtual.
Como pode o conhecimento ser visualizado?	Por meio de arquétipos que contêm representações visuais que "a priori" são conhecidas pelo grupo-alvo.

O motivo para usar arquétipos visuais, contendo representações do mundo real, se baseou por um lado nas habilidades que dos seres humanos em processar imagens rapidamente e por outro lado na facilidade dos indivíduos para relacionar conhecimentos prévios associados a uma imagem já conhecida. Ao ver uma imagem que pertence a sua visão do mundo, as pessoas sabem o que ela representa apesar de não se lembrar dos nomes dos conceitos nela representados.

O modelo proposto é genérico e pode ser usado em qualquer domínio que permita a representação de conceitos em arquétipos. Pode, também, ser usado sobre qualquer repositório de documentos desde que estes possam ser mapeados em representações de conhecimentos, como, por exemplo, ontologias ou taxonomias.

Um dos requerimentos do modelo é a necessidade de mapeamento semântico dos arquétipos e repositórios de documentos. Os repositórios de documentos caracterizam-se por armazenar grandes quantidades de informação fazendo com que o processo de anotação manual seja inviável. Por outro lado, a anotação automática ou semiautomática pode ser facilitada pelos métodos da área da extração de informação, mas a precisão destes métodos pode variar dependendo do domínio e tipo de documento. No caso dos arquétipos visuais, as anotações manuais são viáveis na maioria dos domínios já que a quantidade de conceitos que se espera que sejam representados neles é relativamente pequena em comparação aos conceitos representados nos documentos.

Conclusão

Uma das vantagens que oferece o modelo é a facilidade para recuperação de documentos usando apenas regiões das imagens contidas nos arquétipos. Nos sistemas de RI tradicionais, os usuários traduzem suas necessidades de informação em termos de buscas, estes retornam uma lista de itens que correspondem aos documentos mais relevantes segundo os termos informados. No modelo proposto os usuários traduzem suas necessidades de informação também em termos, mas estes termos não são usados para procurar diretamente os documentos e sim para procurar o arquétipo visual que será usado no processo de visualização.

Outra vantagem do modelo é a integração dos componentes por meio dos mapeamentos semânticos, isto permite que um mesmo arquétipo possa ser usado para visualizar documentos de diferentes repositórios. Por exemplo, o arquétipo do coração apresentado nas seções anteriores, pode ser usado para recuperar artigos científicos, competências organizacionais, projetos, estudos clínicos, indicadores, imagens médicas, entre outros.

Referências

ALAVI, M.; LEIDNER, D. Review: knowledge management and knowledge management systems: conceptual foundations and research issues. **MIS Quarterly**, v. 25, n. 1, p. 107-136, 2001.

BODENREIDER, O. The unified medical language system (UMLS): integrating biomedical terminology. **Nucleic Acids Research**, v. 32, n. 1, p. 267-270, 2004.

BONAMIGO, T. P.; SIQUEIRA, I. Screening for abdominal aortic aneurysms. **Revista do Hospital das Clínicas**, v. 58, n. 2, p. 63-68, 2003.

BURKHARD, R. A. Towards a framework and a model for knowledge visualization: synergies between information and knowledge visualization. In: TERGAN, S. O.; KELLER, T. (Ed.). **Knowledge and information visualization: searching for synergies**. Berlin: Springer, 2005. p. 238-255. (Lecture Notes in Computer Science).

BURKHARD, R.; MEIER, M. Tube map visualization: evaluation of a novel knowledge visualization application for the transfer of knowledge in long-term projects. **Journal of Universal Computer Science**, v. 11, n. 4, p. 473-494, 2005.

CORCHO, O. Ontology based document annotation: trends and open research problems. **International Journal of Metadata, Semantics and Ontologies**, v. 1, n. 1, p. 47-57, 2006.

EPPLER, M.; BURKHARD, R. Visual representations in knowledge management: framework and cases. **Journal of Knowledge Management**, v. 11, n. 4, p. 112-122, 2007.

FEKETE, J. D.; WIJK, J. J.; STASKO, J. T.; NORTH, C. The value of information visualization. In: KERREN, A.; STASKO, J. T.; FEKETE, J. D.; NORTH, C. (Ed.). **Information visualization: human-centered issues and perspectives**. Berlin/Heidelberg: Springer, 2008. p. 1-18. (Lecture Notes in Computer Science).

GERSHON, N.; EICK, S. G.; CARD, S. Information visualization. **Interactions**, v. 5, n. 2, p. 9-15, 1998.

HANBURY, A. A survey of methods for image annotation. **Journal of Visual Languages and Computing**, v. 19, n. 5, p. 617-627, 2008.

HOU, J.; PAI, S. A spatial knowledge sharing platform. Using the visualization approach. **International Journal of Production Research**, v. 47, n. 1, p. 25-50, 2009.

HSU, W.; ANTANI, S.; LONG, L. R.; NEVE, L.; THOMA, G. R. SPIRS: a web-based image retrieval system for large biomedical databases. **International Journal of Medical Informatics**, v. 78, n. 1, p. 13-24, 2009.

KIRYAKOV, A.; POPOV, B.; TERZIEV, I.; MANOV, D.; OGNJANOFF, D. Semantic annotation, indexing, and retrieval. **Web Semantics: Science, Services and Agents on the World Wide Web**, v. 2, n. 1, p. 49-79, 2004.

MÜLLER, H.; MICHOUX, N.; BANDON, D.; GEISSBUHLER, A. A review of content-based image retrieval systems in medical applications--clinical benefits and future directions. **International Journal of Medical Informatics**, v. 73, n. 1, p. 1-23, 2004.

ROSSE, C.; MEJINO, J. L. V. A reference ontology for biomedical informatics: the Foundational Model of Anatomy. **Journal of Biomedical Informatics**, v. 36, n. 6, p. 478-500, 2003.

ROSSE, C.; MEJINO, J. L. V. The foundational model of anatomy ontology. In: BURGER, A.; DAVIDSON, D.; BALDOCK, R. (Ed.). **Anatomy ontologies for**

bioinformatics: principles and practice. London: Springer, 2008. v. 6, p. 59-117. (Computational Biology).
RUBIN, D. L.; MONGKOLWAT, P.; KLEPER, V.; SUPEKAR, K.; CHANNIN, D. S. Annotation and image markup: accessing and interoperating with the semantic content in medical imaging. **Intelligent Systems, IEEE**, v. 24, n. 1, p. 57-65, 2009.
SCHREIBER, A. T.; DUBBELDAM, B.; WIELEMAKER, J.; WIELINGA, B. Ontology-based photo annotation. **Intelligent Systems, IEEE**, v. 16, n. 3, p. 66-74, 2001.
UREN, V.; CIMIANO, P.; IRIA, J.; HANDSCHUH, S.; VARGAS-VERA, M.; MOTTA, E.; CIRAVEGNA, F. Semantic annotation for knowledge management:

Requirements and a survey of the state of the art. **Web Semantics: Science, Services and Agents on the World Wide Web**, v. 4, n. 1, p. 14-28, 2006.

WARE, C. **Information visualization:** perception for design. San Francisco: Morgan Kaufmann Publishers Inc., 2000.

Received on July 19, 2010.

Accepted on June 21, 2011.

License information: This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.