

Representações visuais para recuperação de informação na BDTD- UFPE

Renato Fernandes Correa

Doutor em Ciência da Computação pela UFPE desde 2008 Professor do Departamento de Ciência da Informação e no Programa de Pós-Graduação em Ciência da Informação da UFPE

Jessica Monique de Lira Vieira

Bacharel em Biblioteconomia pela UFPE desde 2011 Mestranda no Programa de Pós-Graduação em Ciência da Informação em 2012

Analisa a aplicabilidade de visualizações, como interfaces amigáveis para Sistemas de Recuperação de Informação (SRIs), bem como discute a construção de representação visual de informações para o SRI da Biblioteca Digital de Teses e Dissertações (BDTD) da Universidade Federal de Pernambuco (UFPE). A metodologia consiste em estudo de caso, envolvendo análise do uso de visualizações em sites de SRIs e discussão dos caminhos percorridos para construção de visualizações para recuperação de teses e dissertações do Programa de Pós-Graduação em Direto da UFPE. Conclui-se que interfaces de SRIs, que disponibilizam visualizações, transmitem informações de forma agradável e interativa, possibilitando a recuperação de documentos relevantes e garantido melhor socialização dos conhecimentos. Entretanto, a escolha da visualização a ser construída para determinado SRI, deve levar em conta não só as vantagens do uso das mesmas pelo usuário, mas, também, o atendimento aos requisitos de organização da informação necessários para construção de cada visualização.

Palavras-Chave: *Sistemas de recuperação de informação; Visualização da informação; Organização da informação; BDTD; Teses e dissertações.*

Visual representations to information retrieval in BDTD-UFPE

This article analyzes the applicability of visualizations as friendly interfaces for Information Retrieval Systems (IRSs) and discusses the construction of visual representation of information to the IRS of the Digital Library of Theses and Dissertations (BDTD) at Federal University of Pernambuco (UFPE). The methodology consists of case study involving analysis of the use of visualizations in SRI sites, and discussion of the paths chosen to build visualizations for retrieval of theses and dissertations of the Law Graduate Program at UFPE. We conclude that SRI interfaces that provide visualizations convey information so enjoyable and interactive, allowing the retrieval of relevant documents and ensuring better socialization of knowledge. However, the choice of one visualization to be built for a specific SRI, should take into account not only the advantages for users of using the same, but also meeting the requirements of information organizing required for construction of each visualization.

Keywords: *Information retrieval systems; Information visualization; Information organization; BDTD; Theses and dissertations.*

Recebido em 04.05.2012 Aceito em 28.11.2013

1 Introdução

Os Sistemas de Recuperação de Informação (SRIs), no contexto deste trabalho, são *softwares* que visam satisfazer às necessidades informacionais dos usuários, auxiliando-os a recuperar conteúdos pertinentes, em meio ao grande volume de informações disponíveis em coleções de documentos (BAEZA-YATES; RIBEIRO NETO, 1999).

O excesso de informações é um dos principais fatores que dificultam a recuperação de documentos relevantes à necessidade do usuário. Além da sobrecarga de informações, os SRIs enfrentam outros problemas, como a dificuldade do usuário em expressar ao sistema sua real necessidade, assim como a dificuldade em dominar o vocabulário usado pelo sistema na indexação, isto é, como estão representadas as informações; e, por outro lado, existe a dificuldade do sistema em compreender a necessidade expressa e recuperar para o usuário os documentos que satisfaçam essa necessidade (SOUZA, 2006).

Tais barreiras poderiam ser minimizadas se os SRIs possuísssem uma interface gráfica amigável, que permitisse uma interação efetiva do usuário com representações gráficas das informações disponibilizadas pelo sistema: facilitando a busca, navegação, localização e recuperação da informação; viabilizando a possibilidade de compreensão de como está organizada a informação no sistema; oferecendo uma visão geral dos assuntos tratados; e contextualizando o resultado da busca, de forma a satisfazer a necessidade informacional do usuário.

A representação das informações de forma gráfica, como afirma Valiati (2008, p. 14), “busca não apenas divulgar ou apresentar dados e resultados, mas prover mecanismos que possibilitem ao usuário interagir com a informação, explorando-a visualmente”. O uso de interfaces de visualização para disponibilização de informações contribui para melhor compreensão pelo usuário e para dedução de novas informações, pois uma informação apresentada de forma gráfica traz consigo um apelo de maior atenção por parte do consumidor do que uma apresentação puramente textual (DIAS; CARVALHO, 2007).

Segundo Estivalet (2000), interfaces visuais permitem observar, manipular, pesquisar, navegar, explorar, filtrar, descobrir, entender e interagir com grandes volumes de dados, muito mais rapidamente e muito mais efetivamente para descobrir padrões ocultos.

Sob esse contexto, o objetivo deste artigo é analisar a aplicabilidade de visualizações como interfaces amigáveis para Sistemas de Recuperação de Informação (SRIs), bem como discutir a construção de representação visual de informações para o SRI da Biblioteca Digital de Teses e Dissertações (BDTD) da Universidade Federal de Pernambuco (UFPE). As visualizações serão analisadas quanto às informações disponibilizadas, vantagens para o usuário e requisitos de organização da informação na construção das mesmas.

2 Metodologia

De acordo com os objetivos, trata-se de uma pesquisa de caráter exploratório, com abordagem qualitativa. Segundo os procedimentos metodológicos, a pesquisa se caracteriza como estudo de caso (GIL, 2010), no qual, primeiramente, analisa-se o uso de visualizações em SRIs e, depois, discute-se a construção de visualização para recuperação das teses e dissertações do Programa de Pós-Graduação em Direito da UFPE, depositadas na BDTD-UFPE.

Inicialmente, foi feito um levantamento de SRIs na *Web*, que fizessem uso de visualizações na interface com o usuário, bem como a identificação de qual técnica de Visualização da Informação ou modelo de interface da área de Ciência da Informação foi utilizado na concepção de cada visualização. Em meio às diversas técnicas de Visualização da Informação existentes, foram detectados SRIs que fizessem uso das seguintes técnicas: desenho de grafos – utilizada no SRI da Queens

Library¹; desenho de árvores – utilizada no mecanismo de busca Yippi²; e *browser* hiperbólico – utilizada no SRI da Infoteca da Embrapa³. Em meio aos modelos de interface discutidos na literatura de Ciência da Informação, encontramos SRIs que fizessem uso dos seguintes modelos: cartografia temática – utilizado no WEBSOM⁴; nuvem de *tags* – utilizado no BRAPCI⁵; e mapas conceituais – utilizado no CMAP TOOLS⁶. Assim, consideramos, neste trabalho, tais técnicas e modelos de interface, como os que mais se adéquam ao contexto dos Sistemas de Recuperação de Informação. Uma avaliação das visualizações quanto aos indicadores propostos pela área de Visualização da Informação pode ser encontrada em (VIEIRA; CORREA, 2011).

Nesse artigo, para cada visualização utilizada nos SRIs, foi realizada uma análise quanto às informações representadas graficamente, as vantagens para o usuário do SRI e requisitos de organização da informação necessários na construção das mesmas.

Por fim, com base no resultado das análises, discute-se o estudo de caso da construção de visualização para a BDTD-UFPE, buscando construir visualizações que permitam recuperar teses e dissertações do Programa de Pós-Graduação em Direito. O objetivo deste estudo de caso é validar os critérios levantados para escolha da construção de visualizações para SRIs, bem como demonstrar a contribuição da visualização no processo de busca e recuperação de informação. Para construção das visualizações, foi adotado um vocabulário controlado, construído com base nos metadados das teses e dissertações do Programa de Pós-Graduação em Direito. Através desse vocabulário controlado, foi possível medir a frequência de coocorrência dos termos nas teses e dissertações, isto é, em quantos documentos cada par de termos aparece como descritores do mesmo documento.

3 Representação visual de informações

Nesta seção, será abordada a área de Visualização da Informação e algumas de suas técnicas, além de modelos de interface discutidos na área de Ciência da Informação. Embora as técnicas de Visualização da Informação e os modelos de interface da Ciência da Informação possam ser utilizados na construção de visualizações para SRIs, são tratados diferentemente neste artigo, devido à gênese em diferentes áreas do conhecimento e à natureza mais geral de aplicação das técnicas, que podem ser utilizadas para construção de visualizações de qualquer domínio, do que dos modelos de interface, os quais se enquadram na construção de visualizações no domínio dos sistemas de informação.

Segundo Card *et al.* (1999 *apud* PRESSER; SILVA; SANTOS, 2010), a Visualização da Informação procura reunir o poder da percepção visual humana e o poder de

¹ Disponível em: <<http://aqua.queenslibrary.org/>>. Acesso em: 11 nov. 2012.

² Disponível em: <<http://search.yippy.com/>>. Acesso em: 11 nov. 2012.

³ Disponível em: <<http://www.infoteca.cnptia.embrapa.br/hypertree/hypertree.html>>. Acesso em: 16 nov. 2012.

⁴ Disponível em: <<http://websom.hut.fi/websom/milliondemo/html/root.html>>. Acesso em: 20 nov. 2012.

⁵ Disponível em: <<http://www.brapci.ufpr.br/>>. Acesso em: 24 nov. 2012.

⁶ Disponível em: <<http://cmap.ihmc.us/conceptmap.html>>. Acesso em: 29 nov. 2012.

processamento do computador para analisar e compreender um conjunto de informações, de forma a otimizar o uso da capacidade visual humana para compreensão das informações transmitidas.

A Visualização da Informação pode ser definida como uma área da Ciência que tem por objetivo o estudo das principais formas de representações gráficas para apresentação de informações, de maneira a contribuir para uma melhor percepção e entendimento das mesmas pelos usuários, além de possibilitar a dedução de novos conhecimentos com base no que está sendo apresentado (FREITAS et al., 2001). Consiste em uma área que combina aspectos de computação gráfica, interação humano-computador, cartografia e mineração de dados. O processo de visualização da informação está relacionado com a transformação de dados abstratos em imagens ou gráficos denominados visualizações.

Como afirma Torres Pojuan e Ponjuan Dante (2010), a Visualização da Informação busca otimizar a carga cognitiva na representação visual das estruturas, relações e padrões de informação e oferecer novas visões analíticas diferentes, a partir dos dados contidos na informação que se representa.

O objetivo na construção das visualizações é auxiliar no entendimento de determinado assunto, minimizando o esforço cognitivo nessa compreensão, visando à inclusão informacional dos usuários ou consumidores das visualizações. Uma simples visualização pode condensar uma grande quantidade de informações, facilitando a compreensão das mesmas, já que a visão é o sentido humano que possui maior capacidade de captação de informações por unidade de tempo (NASCIMENTO; FERREIRA, 2005).

Na maioria dos casos, o uso de recursos gráficos na apresentação de informações produz a compreensão da mensagem transmitida de forma mais natural e com menos esforço cognitivo.

Ao representar graficamente os dados de um determinado domínio de aplicação, de modo que a representação visual gerada explore a capacidade de percepção do homem, permite-se que o usuário, a partir das relações espaciais exibidas, interprete e compreenda as informações apresentadas e, finalmente, deduza novos conhecimentos (FREITAS et al., 2001).

As cinco vantagens que a visualização de informações propicia quando utilizada de forma eficiente, segundo Ware (2000 *apud* LUNARDI, 2008), são: compreensão - a visualização permite a compreensão de grande quantidade de informação; percepção - a visualização revela propriedades do dado que não podem ser antecipadas; controle de qualidade - a visualização permite o controle de qualidade dos dados porque os problemas se tornam imediatamente aparentes; foco mais contexto - a visualização facilita a compreensão de um aspecto dentro do contexto geral dos dados no qual esse encontra; e interpretação - a visualização apóia a formação de hipóteses que propiciam futuras investigações.

3.1 Técnicas de visualização

Será feita aqui uma breve explanação de algumas das diversas técnicas de visualização existentes, sendo abordadas três técnicas atualmente utilizadas em interfaces de SRIs. Maiores detalhes sobre as técnicas de Visualização da Informação podem ser encontrados em (VIEIRA; CORRÊA, 2011).

Na técnica *Desenho de Grafos*, adota-se uma representação gráfica, consistindo em um conjunto de vértices e um conjunto de arestas com localização definida no plano, nos quais os vértices geralmente representam objetos concretos ou abstratos em diversas áreas do conhecimento humano, e arestas indicam relações entre esses objetos. No contexto dos SRIs, os vértices podem representar termos e as arestas relações semânticas ou de coocorrência entre termos.

A técnica *Desenho de Árvores* consiste em uma particularização do desenho de grafos e se destaca pela popularidade e simplicidade de sua estrutura (NASCIMENTO; FERREIRA, 2005). As árvores representam hierarquias, tais como diagramas organizacionais, árvores de diretórios e taxonomias. Os conceitos ou categorias constituem nodos na árvore e se relacionam hierarquicamente com outros conceitos subordinados ou mais específicos por meio de vértices. Os conceitos subordinados só são visualizados caso seja solicitado o detalhamento do conceito mais geral ao qual estão relacionados.

A técnica *Browser Hiperbólico* faz uso do desenho radial de árvores, onde a estrutura de uma árvore é mapeada em uma representação no plano hiperbólico. A técnica é conhecida por utilizar o princípio do foco mais contexto, cujo objetivo é apresentar uma visão geral dos dados a serem visualizados destacando uma área específica de interesse (foco) através da ampliação suave da mesma, enquanto há uma compactação do resto da imagem ou contexto. Na técnica *Browser Hiperbólico*, os conceitos estão organizados por hierarquias (categorias), tendo um assunto mais geral centralizado e, ao redor deste, conceitos subordinados relacionados por meio de vértices em uma estrutura de árvore. Esta técnica consegue disponibilizar cerca de dez vezes mais vértices de uma árvore do que utilizando uma visualização no plano cartesiano, a navegação também se torna mais efetiva e pode ser realizada através de simples movimentos do *mouse*, selecionando o vértice de maior interesse e arrastando-o para o centro para ter o foco desejado (NASCIMENTO; FERREIRA, 2005).

3.2 Modelos de interface

A seguir, serão apresentados três modelos de interface presentes na literatura de Ciência da Informação, que são utilizados em interfaces gráficas de SRIs.

Os *Mapas Conceituais* são modelos gráficos usados para organizar e representar o conhecimento a partir de diagramas que indicam relações

entre conceitos. Nos mapas conceituais, círculos ou retângulos representam conceitos e os relacionamentos entre os conceitos ou as proposições são indicados por uma linha que conecta os dois conceitos, as relações existentes entre eles são especificadas por palavras sobre as linhas (SILVA, 2006). Novak e Cañas (2006) apontam duas importantes características do mapa conceitual: a representação dos conceitos de forma hierárquica, com o conceito mais geral no início do mapa e, depois, os mais específicos, arranjados hierarquicamente; a existência de referências cruzadas que permitem verificar como é representada a relação dos conceitos no domínio do conhecimento. Desta forma, a representação gráfica e concisa do conhecimento, através do mapa conceitual e a possibilidade de ligação entre conceitos, representados por termos ou palavras-chave, resultam em um ambiente ideal para se definir um sistema de navegação, no qual os usuários podem encontrar a informação que buscam, bem como navegar pelo sistema pesquisando outros temas.

A *Cartografia Temática* é uma técnica bibliométrica avançada, consistindo em um mecanismo de representação das temáticas inscritas em recursos informacionais, tais como livros, periódicos, teses, etc. (ELIEL, 2007). Pode ser entendida como mapas gerados a partir de dados bibliográficos, proporcionando uma visualização gráfica das temáticas tratadas, oferecendo uma visão global de conjuntos de informações e evidenciando relações e estruturas entre elas. Esta técnica facilita a compreensão das relações existentes entre os dados, quando os mesmos são expressos por meio de mapas ou cartografias, e configura-se como interface poderosa para uso em Sistemas de Recuperação de Informação (KOBASHI; SANTOS, 2008), pois permite a exploração rápida de conjuntos de informações desconhecidas; evidenciação de relações e estruturas nas informações; fornecimento de alternativas de acesso a informações pertinentes; e classificação interativa de informação.

A Nuvem de Tags ou nuvem de etiquetas (do inglês Tag Cloud) consiste em uma interface textual usada por websites e blogs (MOREIRA; ROMÃO, 2009). Na nuvem de tags, as palavras-chave são dispostas em lista visual com o objetivo de representar os itens de conteúdo de um site, mostrando aos usuários quais são os assuntos principais abordados por páginas deste site. A nuvem de tags pode, também, retratar as tags mais acessadas pelos usuários no momento, ou seja, as palavras mais cobiçadas pelos usuários. O tamanho da fonte de cada etiqueta representa proporcionalmente a quantidade de conteúdos que o site apresenta sobre o assunto ou o número de acesso àquela etiqueta. As tags que estiverem em destaque dentro da nuvem são as de maior ocorrência no momento, representando, assim, os assuntos e ou informações que estão em foco dentro do site. Cada etiqueta disponibilizada na nuvem funciona como link que leva à coleção de documentos relacionados ou indexados com aquela etiqueta. A nuvem de tags facilita a navegação, recuperação e compartilhamento de informações pelo usuário-navegador, visto que sua estrutura, como componente de busca dentro do site, situa os assuntos mais populares na indexação de documentos através de uma indicação

visual amigável, possibilitando ao usuário que visita o site estar situado sobre as novidades visualizando as tags em evidência.

4 Resultados

Nesta seção, será apresentado o resultado da análise da aplicação de visualizações em SRIs, bem como o estudo de caso de construção da visualização para o Programa de Pós-Graduação em Direito.

4.1 Resultados da análise das visualizações utilizadas em SRI

O Quadro 1 traz para cada tipo de visualização (técnica de Visualização da Informação ou modelo de interface utilizado na concepção da visualização): as informações representadas graficamente, as vantagens do uso da visualização para os usuários dos SRIs e os requisitos de organização da informação necessários para a construção da visualização.

De acordo com o Quadro 1, as visualizações baseadas no desenho de grafo, árvores e mapas conceituais exigem a utilização de instrumentos de organização da informação que demandam grande esforço intelectual e tempo para serem construídos, como: taxonomias, tesouros e mapas conceituais. Já as visualizações baseadas em nuvem de *tags* e cartográfica temática requerem a utilização de listas de cabeçalho de assuntos, visando à normalização e padronização dos termos. A lista de cabeçalho de assunto é um instrumento de organização da informação que exige menos esforço intelectual e tempo para elaboração.

Quanto às informações representadas graficamente, as visualizações são utilizadas principalmente para apresentar o espaço dos termos ou vocabulário do sistema aos usuários. Somente a cartográfica temática e o desenho de árvores têm sido utilizados para apresentar, além o espaço de termos, o espaço de documentos, buscando auxiliar o usuário do SRI a entender a relação de similaridade entre os documentos.

Na maioria dos casos, ao revelar o vocabulário do sistema, as visualizações trazem para o usuário do SRI a vantagem de permitir a especificação da busca ou refinamento da estratégia de busca. Ao representar o espaço de documentos, as visualizações baseadas em cartografia temática e desenho de árvores, podem facilitar a discriminação dos documentos constantes no resultado de uma busca

Quadro 1 – Critérios para escolha de visualizações por SRIs

Tipo de Visualização	Informações representadas	Vantagens para o usuário	Requisitos
Desenho de Grafos	Termos relacionados a determinado termo; diferentes tipos de relações entre termos, onde termos são representados por meio de vértices e relações entre eles por meio de arestas.	Facilita o refinamento da busca; permite a localização de termos relacionados ao que foi pesquisado.	Adoção de um tesauro ou taxonomia, que defina as relações existentes entre os termos.
Desenho de Árvores	Termos tratados e relação de subordinação entre termos dispostos na forma de árvore; visualização dos documentos relacionados a cada termo.	Possibilita especificar uma busca ao selecionar termos; descobrir relações de subordinação entre termos; ao representar dinamicamente os resultados da busca, minimiza a sobrecarga de informação e facilita a discriminação da informação, reunindo documentos em grupos.	Adoção de um tesauro ou taxonomia que deixe claro as relações hierárquicas existentes entre os termos, bem como entre termos e documentos.
Browser Hiperbólico	Assuntos tratados e relações de subordinação entre assuntos, representados numa árvore no plano hiperbólico.	Possibilita visualizar, de forma gráfica, como estão estruturas as informações; facilita a especificação da busca.	Adoção de uma taxonomia que defina a hierarquia dos termos e a associação entre termos e documentos.
Mapa Conceitual	Conceitos tratados; preposições verdadeiras entre conceitos, representados como mapas conceituais.	Disponibiliza o conhecimento a ser recuperado, explicitando como todo o conteúdo semântico é organizado e como se conecta internamente; visualiza documentos relacionados a cada conceito.	Construção de um mapa conceitual que permita visualizar os conceitos e as relações estabelecidas entre conceitos; associação dos documentos aos conceitos.
Cartografia Temática	Assuntos mais comuns e relação de coocorrência entre assuntos; relação de similaridade entre documentos.	Proporciona uma visão global dos assuntos tratados; e evidencia relações de coocorrência entre eles; permite localizar documentos similares em conteúdo.	Adoção de um vocabulário controlado que permita mensurar a coocorrência entre termos e a relação de similaridade entre documentos.
Nuvem de Tags	Assuntos mais abordados ou assuntos mais consultados representados por tags; as <i>tags</i> são formatadas de acordo com a frequência relativa e dispostas em ordem alfabética em uma nuvem.	Serve como sugestão de busca para os usuários permitindo explorar a coleção de documentos através dos assuntos mais tratados ou mais procurados; disposição alfabética dos assuntos; permite o refinamento da expressão de busca, se construída dinamicamente a partir dos resultados da busca.	Adoção de uma lista de termos de indexação que sirva de base pra organização das tags; aferir a frequência de ocorrência de cada <i>tag</i> e a associação entre <i>tags</i> e documentos.

Fonte: Elaborado pelo autor.

Os critérios levantados no Quadro 1 podem servir de guia para a escolha de construção de visualizações para SRIs em outros contextos (domínios e áreas de conhecimento).

4.2 Estudo de caso

A seguir, será apresentado o estudo de caso da construção de visualização para a BDTD-UFPE, que visa a validar os critérios levantados para escolha da construção de visualizações para SRIs, bem como demonstrar a contribuição da visualização no processo de busca e recuperação de informação.

4.2.1 Escolha do tipo de visualização para BDTD-UFPE

A BDTD-UFPE (Biblioteca Digital de Teses e Dissertações da Universidade Federal de Pernambuco) tem como objetivo disponibilizar, na íntegra, as teses e dissertações produzidas pelos programas de pós-graduação da instituição. A BDTD adota o Padrão Brasileiro de Metadados para Teses e Dissertações (MTD-BR) para descrever eletronicamente os metadados das teses e dissertações, visando ao intercâmbio entre sistemas.

O atual Sistema de Recuperação de Informação da BDTD-UFPE, parte integrante do sistema TEDE disponibilizado pelo IBICT, somente dá aos usuários a possibilidade de explorar a coleção de documentos por meio de buscas e navegação por teses e dissertações defendidas em cada programa de pós-graduação. Essa metodologia de pesquisa oferece sérias limitações, devido à dificuldade dos usuários em expressarem o que procuram através de uma consulta e em conhecer a linguagem usada pelo sistema para representar os tópicos presentes nas teses e dissertações. Essas limitações constituem-se entraves à acessibilidade local e nacional da informação científica e tecnológica produzida nos programas de pós-graduação na UFPE.

No SRI da BDTD-UFPE não há uma representação visual das informações contidas nos documentos. Os documentos estão estruturados por temática, de acordo com o número de classificação na CDD, a partir do qual é possível compreender de que se trata o documento. Também é possível compreender de que se trata o documento através do valor do campo assunto, selecionado pelo autor da Tese ou Dissertação, através da tabela de área de conhecimento do CNPq, geralmente a grande área do Programa de Pós-Graduação.

Entretanto, essas temáticas são muito gerais, sendo necessária, para contemplar assuntos mais específicos, a construção de taxonomias e/ou tesouros, que demandam um tempo considerável e grande esforço intelectual para construção.

Modelos de interface que poderiam ser aplicados na BDTD-UFPE, em curto prazo, são a cartografia temática e a nuvem de *tags*, visto que são técnicas que não necessitam da organização dos termos ou conceitos em hierarquia, sendo necessário apenas a padronização ou normalização dos termos.

A aplicação da cartografia temática facilitaria a recuperação de informações, pois a adoção da interface de navegação amigável, que possibilite ao usuário uma melhor visualização de como estão organizados os termos, evidenciando as relações de coocorrência estabelecidas entre eles, proporciona uma melhor localização da informação desejada.

A adoção de uma interface de navegação baseada em nuvem de *tags* é, também, uma maneira de tornar o SRI da BDTD-UFPE mais amigável, sendo capaz de revelar visualmente os assuntos, a frequência com que ocorrem e os respectivos documentos que os abordam dentro do repositório.

O Quadro 2 apresenta, de maneira resumida, as visualizações escolhidas que poderiam ser utilizadas pela BDTD-UFPE e que vantagens trariam para os usuários.

Assim, serão construídas e analisadas visualizações baseadas nos modelos de interface 'cartografia temática' e 'nuvem de tags' para teses e dissertações do Programa de Pós-Graduação em Direito depositadas na BDTD-UFPE, com o objetivo de potencializar a interação do usuário com a biblioteca digital, minimizando, dessa forma, as barreiras que dificultam a acessibilidade.

Quadro 2 – Visualizações para BDTD-UFPE, requisitos para organização da informação e vantagens oferecidas ao usuário

Tipo de Visualização	Requisitos na organização da informação da BDTD-UFPE	Vantagens para o usuário
Cartografia Temática	Adoção de uma lista de cabeçalho de assuntos. Não necessitam da organização dos termos ou conceitos por hierarquia, sendo necessário um instrumento para padronização ou normalização dos termos.	Possibilita melhor visualização de como estão organizados os termos e a coocorrência estabelecidas entre eles; proporciona localização da informação desejada por meio da navegação.
Nuvem de Tags		Revela visualmente os tópicos, a frequência com que ocorrem e os respectivos documentos que abordam cada tópico dentro do repositório; facilita a recuperação dos documentos a partir da seleção de uma tag.

Fonte: Elaborado pelo autor.

4.2.2 Construção e análise de visualizações para as teses e dissertações de Direito

Para auxiliar a construção de uma visualização para o Programa de Pós-Graduação em Direito, foi adotado, neste trabalho, um vocabulário controlado elaborado por Braz (2010), que consiste em uma lista de cabeçalhos de assunto composta por 145 termos, dos quais 122 entradas principais e mais 23 entradas remissivas. A metodologia de construção do vocabulário é descrita em (BRAZ; CORREA, 2011). Através desse vocabulário controlado, foram construídas duas visualizações que poderiam ser utilizadas pelo Programa de Pós-Graduação em Direito da BDTD-UFPE.

A visualização da Figura 1 foi construída, adotando-se a nuvem de tags, nesta visualização estão presentes os termos (tags) que representam as principais temáticas das teses e dissertações depositadas no programa, o tamanho da fonte das tags representa proporcionalmente a quantidade de documentos que tratam de cada uma, a quantidade exata de documentos em cada tag pode ser percebida ao mover o mouse sobre elas. As tags se encontram em ordem alfabética, o que facilita a localização dos assuntos. Ao clicar sobre uma tag, os documentos a ela associados seriam listados.

A visualização baseada na nuvem de tags não evidencia relações hierárquicas ou de proximidade entre os termos, o que pode representar uma limitação para o usuário, no entanto, o fato de organizar os termos

em ordem alfabética facilita a localização e a recuperação das teses e dissertações sobre cada assunto.

As visualizações apresentadas nas Figuras 2 e 3 foram construídas por meio da cartografia temática.

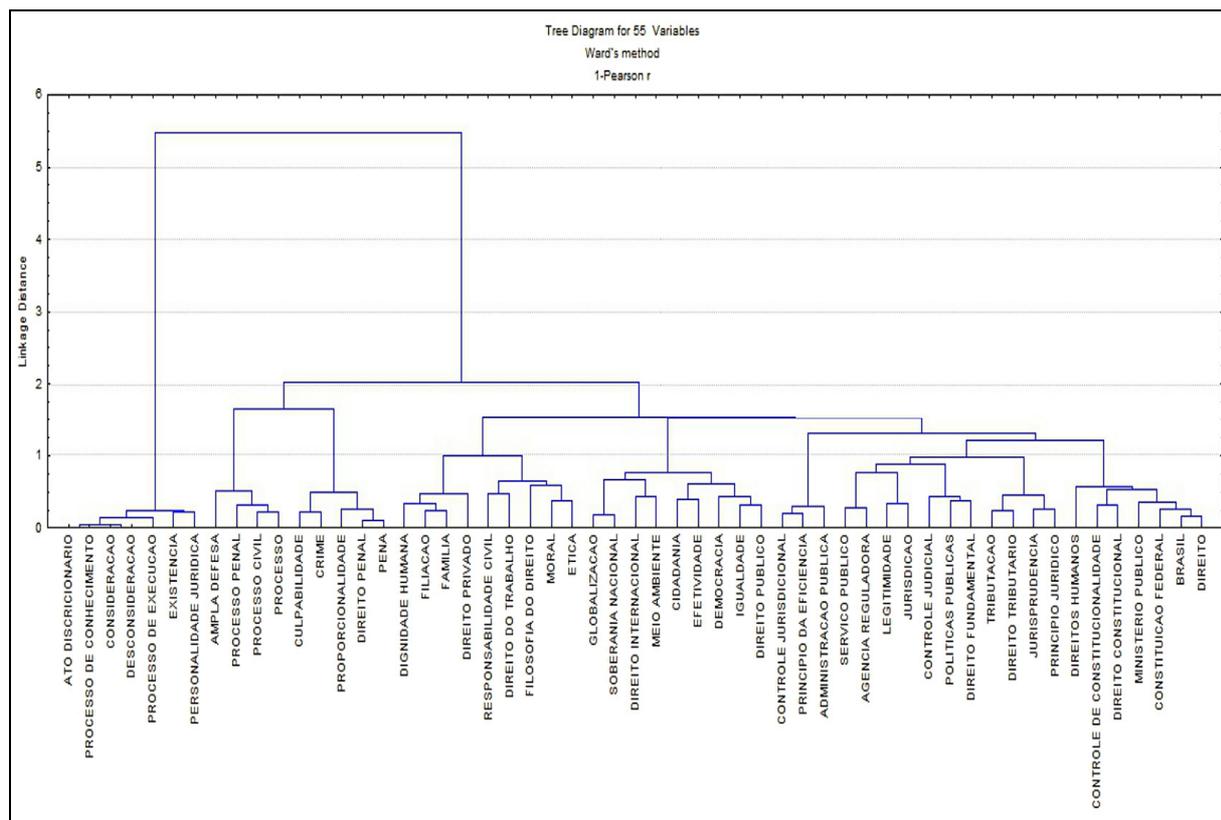
Figura 1 – Nuvem de Tags construída para o Programa de Pós-Graduação em Direito



Fonte: Dados da pesquisa.

A Figura 2 traz a classificação hierárquica de descritores em um dendograma, permitindo visualizar as relações de proximidade temática entre os termos, aferida da relação de coocorrência dos descritores na indexação de teses e dissertações. Na figura 2, são apresentados somente os descritores com frequência maior ou igual a 20.

Figura 2 – Cartografia temática para o Programa de Pós-Graduação em Direito, baseada em classificação hierárquica de descritores-
DENDOGRAMA

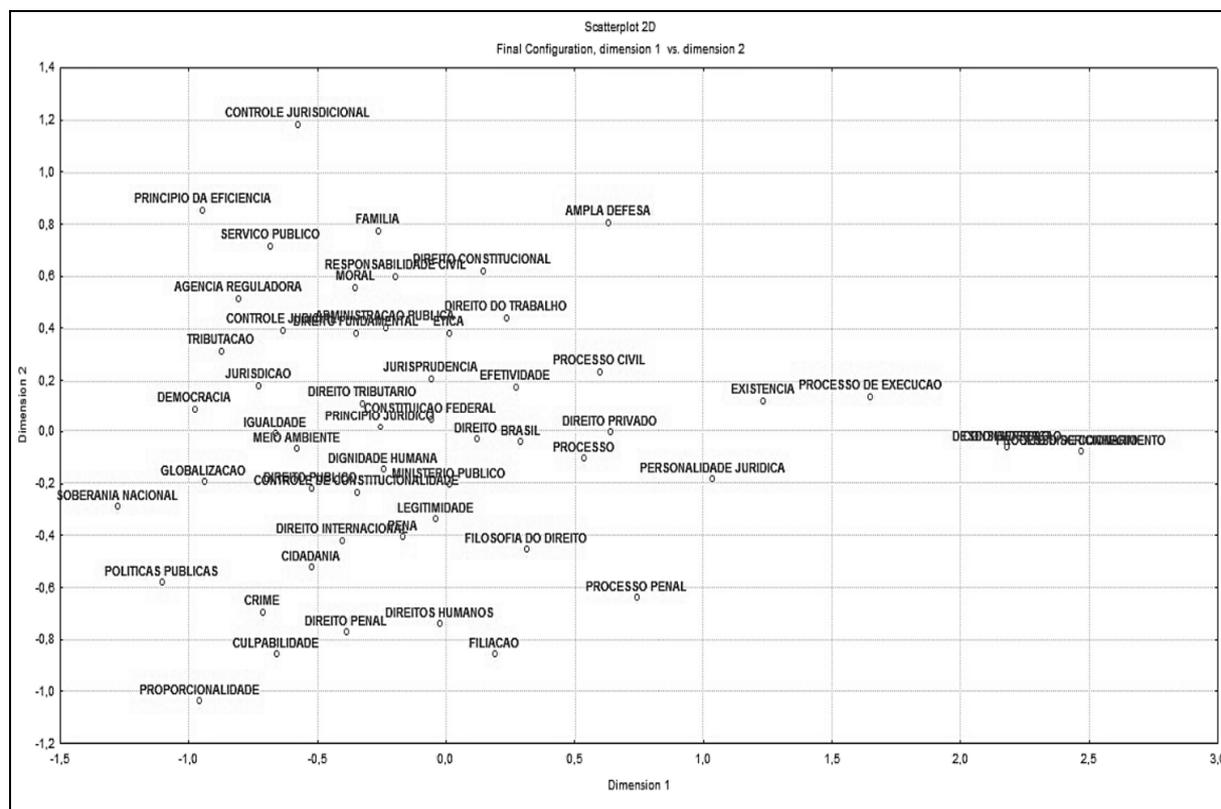


Fonte: Dados da pesquisa.

A visualização da Figura 2 apresenta os descritores em uma hierarquia criada a partir das relações de coocorrência entre os termos, essa visualização apresenta como limitação a sobrecarga de informações e pode levar a erros de interpretação, já que a relação de coocorrência não implica na relação de subordinação entre os termos.

A Figura 3 traz a visualização dos descritores em espaço reduzido, criada através da escala multidimensional (MDS) sobre a matriz de coocorrência dos termos. Apenas os descritores com frequência maior ou igual a 20 aparecem no mapa. A proximidade dos termos no mapa indica a relação de associação entre eles, isto é, os conceitos mais próximos no mapa são os que possuem uma proximidade temática maior ou maior coocorrência.

Figura 3 – Cartografia temática para o Programa de Pós-Graduação em Direito, baseada em ESCALA MULTIDIMENSIONAL (MDS)



Fonte: Dados da pesquisa.

A visualização da Figura 3 apresenta uma estrutura não hierárquica dos descritores, mas representa graficamente a proximidade entre as temáticas. A visualização também gera uma sobrecarga de informações. As relações de coocorrência permitem a visualização das relações associativas entre as temáticas das teses e dissertações do programa dentro de uma estrutura gráfica, de maneira mais interativa do que em estruturas não gráficas ou vocabulários controlados por si só.

Tanto a visualização baseada na nuvem de *tags* (figura 1) quanto à visualização baseada na cartografia temática (Figura 2 e Figura 3) revelam, de forma gráfica, os termos que representam as temáticas das teses e dissertações do programa de Pós-Graduação em Direito. Entretanto, a nuvem de *tags* não revela relações entre os termos, enquanto a visualização por cartografia temática evidencia as relações de coocorrência entre as temáticas através da proximidade dos termos na visualização.

Comparando as Figuras 2 e 3, percebe-se que a proximidade dos termos depende do algoritmo computacional utilizado para mensurar esta proximidade relativa, por exemplo, o termo Direitos Humanos se encontra mais próximo de Direito Penal na figura 3 e já na figura 2 é mais próximo de Direito Constitucional

5 Conclusão

No presente contexto, em que os Sistemas de Recuperação de Informação têm enfrentado dificuldades em exercer seu objetivo principal, o de recuperar informações relevantes para seus usuários, o uso de visualizações na interface dos SRIs pode torná-los mais eficazes na comunicação com o usuário e na recuperação da informação.

As técnicas de Visualização da Informação e modelos de interfaces visuais, discutidos na literatura da Ciência da Informação, possibilitam a construção de visualizações para SRIs que podem auxiliar na busca, na obtenção de informações e na dedução de novos conhecimentos a partir do que está sendo apresentado.

Ao analisar os requisitos de cada tipo de visualização (técnica ou modelo de interface) quanto à organização da informação, fica claro que a adoção de instrumentos como vocabulários controlados, taxonomias e tesouros é necessária para garantir visualizações de qualidade.

Dentre as técnicas e modelos de interfaces visuais analisadas no decorrer do trabalho e com base também nas visualizações construídas para recuperação de teses e dissertações do Programa de Pós-Graduação em Direito da UFPE, chegou-se a conclusão que o tipo de visualização que melhor se adéqua como interface para esse SRI é a nuvem de *tags*, visto que, possibilita ao usuário uma melhor visualização dos principais assuntos tratados nos documentos, através das *tags* dispostas, oferecendo uma visão global do conjunto de informações, permitindo visualizar a frequência com que ocorrem e os respectivos documentos que abordam cada assunto dentro do repositório.

Na literatura de Ciência da Informação, são poucos os trabalhos que discutem as visualizações como interfaces para SRIs. Tal fato sugere a necessidade de pesquisas que busquem explorar, de maneira mais específica, a aplicação das visualizações na recuperação de informação.

Nesse contexto, vale ressaltar que, além da BDTD, sistemas de recuperação de informação que apresentem suas informações de maneira visual podem auxiliar o usuário a especificar sua busca, possibilitando uma melhor interação com as informações armazenadas e conseqüentemente a recuperação de informações adequadas a sua necessidade.

A pesquisa, até aqui desenvolvida, pode servir de base para trabalhos futuros, que busquem explorar a visualização da informação na construção de interfaces para sistemas de recuperação de informação, levando em conta as especificidades do contexto em que está inserido o usuário, o domínio de conhecimento do mesmo e sua competência informacional.

Referências

BAEZA-YATES, R.; RIBEIRO-NETO, B. *Modern information retrieval*. New York: ACM Press, 1999.

BRAZ, M. I. Vocabulário controlado para indexação de teses e dissertações do Programa de Pós-Graduação em Direito da UFPE. 2010. 52f. Monografia (Bacharelado em Biblioteconomia) – Universidade Federal de Pernambuco, Recife, 2010.

BRAZ, M. I.; CORREA, R. F. Vocabulário controlado para indexação de teses e dissertações do Programa de Pós-Graduação em Direito da UFPE. In: SILVA, F et al. (Org.). A Gestão da Informação na era do conhecimento. Recife: NECTAR, 2011. 1 CD-ROM.

CARD, S. K.; MACKINLAY, J. D.; SCHNEIDERMAN, B. (Ed.). Readings in information visualization: using vision to think. [s.l.]: Morgan Kaufmann, 1999.

DIAS, M. P.; CARVALHO, J. O. F. de. A visualização da informação e a sua contribuição para a Ciência da Informação. *DataGramaZero: Revista de Ciência da Informação*, v. 8, n. 5, out. 2007. Disponível em: <http://dgz.org.br/out07/Art_02.htm>. Acesso em: 28 out. 2009.

ELIEL, O. *Cartografia temática de artigos de periódicos nacionais da Ciência da Informação (período de 1986 a 2005): contribuições ao campo teórico da área*. 2007. 120 f. Dissertação (Mestrado) - Programa de Pós-graduação em Ciência da informação, Pontifícia universidade Católica de Campinas, Campinas, 2007. Disponível em: <http://www.bibliotecadigital.puc-campinas.edu.br/tde_busca/arquivo.php?codArquivo=278>. Acesso em: 24 fev. 2010.

ESTIVALET, L. F. *O uso de ícones na Visualização de Informações*. 2000. 90f. Dissertação (Mestrado em Ciência da Computação) – Universidade Federal do Rio Grande do Sul, Porto Alegre, 2000.

FREITAS, C. M. D. S. et al. Introdução à Visualização de Informações. **RITA** – Revista de Informática Teórica e Aplicada, Instituto de Informática, v. 8, n. 2, p. 143-158, 2001.

GIL, A. C. *Como elaborar projetos de pesquisa*. 5. ed. São Paulo: Atlas, 2010.

KOBASHI, N.Y.; SANTOS, R. N. M. dos. Arqueologia do trabalho imaterial: uma aplicação bibliométrica à análise de dissertações e teses. *Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação*, Florianópolis, n. esp., 1º sem. 2008. Disponível em: <<http://www.periodicos.ufsc.br/index.php/eb/article/view/1130>>. Acesso em: 17 nov. 2009.

LUNARDI, M. S. *Visualização em nuvens de texto como apoio à busca exploratória na web*. 2008. 126f. Dissertação (Mestrado em Design) - Escola Superior de Desenho Industrial, Universidade do Estado do Rio de Janeiro, São Paulo, 2008. Disponível em: <http://www.bdt.d.uerj.br/tde_busca/arquivo.php?codArquivo=1522>. Acesso em: 11 ago. 2011.

MOREIRA, V. L.; ROMÃO, L. M. S. O funcionamento discursivo das nuvens de tags na rede eletrônica: sentidos sobre Capitolina. *DataGramaZero*, v. 10, n. 6, nov./dez. 2009. Disponível em: <http://www.dgz.org.br/dez09/Art_02.htm>. Acesso em: 3 mar. 2010.

NASCIMENTO, H. A. do; FERREIRA, C. B. R. Visualização de informações: uma abordagem prática. In: CONGRESSO DA SOCIEDADE BRASILEIRA DE COMPUTAÇÃO, 25., 2005, UNISINOS. *Anais...* Rio Grande do Sul: São Leopoldo, 2005. p. 1262-1312.

NOVAK, J. D.; CAÑAS, A. J. *The theory underlying concept maps and how to construct them*. Technical Report, Florida Institute for Human and Machine Cognition (IHCM), 2006, 36 p. Disponível em: <<http://cmap.ihmc.us/Publications/ResearchPapers/TheoryUnderlyingConceptMaps.pdf>>. Acesso em: 23 ago. 2011.

PRESSER, N. H.; SILVA, E. L. da; SANTOS, R. N. M. dos. Recursos de formulação e visualização de indicadores para apoiar processos de gestão educacional em IESS. *Revista Digital de Biblioteconomia e Ciência da Informação*, Campinas, v. 7, n. 2, p. 247-259, jan./jun. 2010.

SILVA, M. F. Diagramas hierárquicos, mapas conceituais e mapas hiperbólicos: um estudo comparativo entre interfaces hipertextuais. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 7., 2006, Marília/ SP. *Anais Eletrônicos...* Marília: ANCIB, 2006. 12 p.

SOUZA, R. R. Sistemas de recuperação de informação e mecanismos de busca na web: panorama atual e tendências. *Perspectivas em ciência da informação*. Belo Horizonte, v. 11, n. 2, p. 161-173, maio/ago. 2006.

TORRES PONJUAN, D.; PONJUÁN DANTE, G. Influencias ergonómicas en la Visualización de Información. *Informação & Informação*, Londrina, v. 15, n. 1, p. 17-36, jan./jun. 2010.

VALIATI, E. R. de A. *Avaliação de usabilidade de técnicas de visualização de informações multidimensionais*. 2008. 220f. Tese (Doutorado em Ciência da Computação) - Universidade Federal do Rio grande do Sul. Porto Alegre, 2008.

VIEIRA, J. M. L.; CORREA, R. F. Visualização da Informação na construção de interfaces amigáveis para Sistemas de Recuperação de Informação. *Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação*, Florianópolis, v. 16, n. 32, 2011. Disponível em: <<http://www.periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2011v16n32p73>>. Acesso em: 28 fev. 2012.

WARE, C. *Information visualization: perception for Design*. [s.l.]: Morgan Kaufmann, 2000.