



UM SISTEMA LIVRE PARA AUTOMATIZAÇÃO DO PADRÃO DE DESCRIÇÃO DA INFORMAÇÃO

A FREE SYSTEM AUTOMATION STANDARD DESCRIPTION OF INFORMATION

Elisa Yumi Nakagawa¹

Thiago Bianchi²

Maria Cristina Ferreira de Oliveira³

Elaine Parros M. de Sousa⁴

Gabriel de Faria Andery⁵

José Carlos Maldonado⁶

RESUMO: A organização e a disseminação de conhecimento contido em acervos históricos têm sido uma constante preocupação; entretanto, há uma carência de sistemas de software para esse domínio. Nesse contexto, foi desenvolvido um sistema web, baseado no Padrão de Descrição da Informação, que viabiliza a representação integrada de informações de diferentes tipos de acervos. Este artigo apresenta a experiência no estabelecimento desse padrão e no desenvolvimento desse sistema. Sendo um software livre, espera-se que pesquisadores e a sociedade em geral se beneficiem desse sistema.

PALAVRAS-CHAVE: Acervo histórico; Sistema web; Software livre.

ABSTRACT: *The organization and dissemination of knowledge contained in historical collections have been a constant concern; however, there is a lack of software systems for this domain. In this context, a web system based on the standard description of information that allows for the integrated representation of information from different types of collections was developed. This article presents the experience in setting this standard and developing the system. Being open source, it is expected that researchers and society at large to benefit from this system.*

KEYWORDS: *History Collection; Web system; Free software.*

¹ Professora associada do Instituto de Ciências Matemáticas e de Computação da Universidade de São Paulo – ICMC/USP. E-mail: elisa@icmc.usp.br.

² Doutorando do Instituto de Ciências Matemáticas e de Computação da Universidade de São Paulo – ICMC/USP. E-mail: tbianchi@icmc.usp.br.

³ Professora titular do Instituto de Ciências Matemáticas e de Computação da Universidade de São Paulo – ICMC/USP. E-mail: cristina@icmc.usp.br.

⁴ Professora doutora do Instituto de Ciências Matemáticas e de Computação da Universidade de São Paulo – ICMC/USP. E-mail: parros@icmc.usp.br.

⁵ Graduado pelo Instituto de Ciências Matemáticas e de Computação da Universidade de São Paulo – ICMC/USP. E-mail: gandery@gmail.com.

⁶ Professor titular do Instituto de Ciências Matemáticas e de Computação da Universidade de São Paulo – ICMC/USP. E-mail: jcmaldon@icmc.usp.br.

Recebido em: 15/10/2013 - **Aceito em:** 28/01/2014

1 INTRODUÇÃO

Uma diversidade de acervos de inestimável valor histórico e cultural contendo objetos de diversas naturezas pode ser encontrada espalhada em diversas instituições, como museus, bibliotecas, fundações públicas e privadas, entre outras. Contudo, no Brasil, muitos desses acervos encontram-se em estado precário de conservação e muitos acervos são desconhecidos, até mesmo por pesquisadores da área. Diversas iniciativas têm sido observadas, tanto públicas quanto privadas, no sentido de organizar e disponibilizar o conteúdo de tais acervos. Esses esforços seriam mais eficientes se contassem com o apoio computacional adequado, no entanto, ainda observa-se uma carência de sistemas de software que auxiliem efetivamente o registro e disseminação de informações sobre esses acervos. Recentemente, Beaudoin (2012a) apresentou uma discussão ampla sobre a necessidade e dificuldade de preservação digital de objetos culturais, ou seja, dos acervos históricos. A autora destaca que a preservação digital aborda diversos contextos, tais como o tecnológico e o de utilização, que devem ser considerados e que certamente influenciam no sucesso da atividade de preservação (BEAUDOIN, 2012a).

Nesse cenário, o Projeto Memória Virtual de São Carlos (MEMÓRIA, 2004); (NAKAGAWA; CRISTIANINI; MORAES, 2006a) — um projeto de políticas públicas — abordou a informatização dos acervos históricos da região da cidade de São Carlos, visando viabilizar o acesso, organização e preservação do patrimônio histórico e cultural das fazendas e museus da região, integrando e disponibilizando acervos do município dispersos geograficamente. O projeto, de natureza multidisciplinar, envolveu a Prefeitura Municipal de São Carlos, o Museu de São Carlos⁷, a Fundação Pró-Memória de São Carlos⁸, a Associação Pró-Casa do Pinhal⁹ e outras instituições, bem como pesquisadores de diferentes áreas de especialidade, como Ciências de Computação, Ciência da Informação, Biblioteconomia, Arquitetura e Museologia.

Um dos principais resultados do projeto Memória Virtual de São Carlos (MEMÓRIA, 2004) foi o estabelecimento de um padrão, denominado Padrão de Descrição da Informação (PDI) (CRISTIANINI et al., 2004), que possibilita a descrição integrada de informações de diversos tipos de acervos históricos em uma única estrutura. Buscando viabilizar a automatização desse padrão, um sistema web, denominado Sistema Memória Virtual¹⁰, foi construído para ser utilizado em instituições que possuam algum tipo de acervo histórico, seja bibliográfico, museológico, arquivológico, arquitetônico ou relacionado a bens naturais. Esse sistema é um software livre, e como tal, fornece aos seus usuários as seguintes liberdades (Free

⁷ <http://www.saocarlos.sp.gov.br/index.php/cultura/115301-museu-sao-carlos.html>

⁸ <http://www.promemoria.saocarlos.sp.gov.br/>

⁹ <http://www.casadopinhal.com.br/>

¹⁰ O Sistema Memória Virtual está disponível em <http://memoriavirtual.icmc.usp.br>. Este artigo é baseado na versão 1.0 do sistema. A versão 2.0 encontra-se em fase de evolução e inserção de novas funcionalidades.

Software Foundation - FSF, 2013): (i) executar o programa para qualquer propósito; (ii) estudar o funcionamento do programa e modificá-lo, adaptando-o às suas necessidades; e (iii) redistribuir cópias do código fonte, juntamente com possíveis alterações introduzidas. Assim, diferentemente de software proprietário, um software considerado livre é acompanhado de uma permissão para uso, distribuição e redistribuição, com ou sem modificações, e na maioria dos casos, de forma gratuita.

O objetivo deste artigo é apresentar a experiência no estabelecimento do padrão PDI e no desenvolvimento do Sistema Memória Virtual. Assim, o artigo está organizado da seguinte maneira. Na Seção 2 são apresentados trabalhos relacionados. Na Seção 3 é apresentado o PDI. Na Seção 4 é apresentado o Sistema Memória Virtual e questões relacionadas ao seu desenvolvimento. Na Seção 5 é relatada a utilização desse sistema. Por fim, na Seção 6 são apresentadas as conclusões e possíveis desdobramentos futuros.

2 TRABALHOS RELACIONADOS

Diversas iniciativas de uso de sistemas de software estão em andamento visando ao registro e recuperação de informações sobre objetos contidos em acervos históricos de diversas naturezas, tais como documentos bibliográficos, documentos de arquivo, objetos museológicos, bens arquitetônicos e bens naturais, identificados neste trabalho como bens patrimoniais. Tais iniciativas introduzem a necessidade de criar novos padrões e normas para uma descrição mais adequada desses acervos. A utilização de padrões e normas na descrição de bens patrimoniais preserva a qualidade das informações disponibilizadas, além de possibilitar a integração de várias redes de informação nacionais e/ou internacionais. Nesse contexto, diversos mecanismos tecnológicos têm sido investigados para dar apoio à preservação digital de acervos históricos, tais como o uso de metadados e métodos, alguns dos quais sumarizados no trabalho de Beaudoin (2012a).

No caso de documentos bibliográficos, as normas contidas no *Anglo-American Cataloguing Rules, 2. Edition* (AACR2) (Anglo-American Cataloguing Rules - AACR, 2013) e no formato *MACHine-Readable Cataloging* (MARC) (MARC, 2013), — que permite a leitura por computador de qualquer informação bibliográfica, compreendendo um conjunto de normas para identificação, armazenamento e intercâmbio de dados — mostram-se insuficientes para atender às novas necessidades impostas pela automatização dos sistemas de software que tratam de acervos de documentos bibliográficos. Para a descrição e padronização de documentos de arquivo, existe a *General International Standard Archival Description* (ISAD(G)) (International Council on Archives - ICA, 2013) que estabelece diretrizes gerais para a preparação de descrições arquivísticas. De acordo com a *International Council on Archives* (ICA), esse padrão deve ser usado em conjunto com outras normas existentes, complementando a descrição dos documentos segundo as necessidades e

especificidades de cada instituição ou acervo. Também existem trabalhos voltados para a descrição de acervos museológicos, como as categorias para a descrição de objetos de arte (*Categories for the Description of Works of Art - CDWA*) (Getty, 2013), que apresentam um conjunto bastante completo de atributos que possibilitam representar diferentes informações associadas a objetos de arte e cultura. Além disso, o CDWA possui um esquema em XML (*eXtensible Markup Language*) (RAY, 2003) que possibilita o intercâmbio de informações de objetos de artes descritos utilizando o CDWA com outros formatos, tais como o MARC.

Paralelamente, existem diversas iniciativas de automatização das tarefas de armazenamento e recuperação de informações relacionadas aos acervos históricos. Pode-se citar, a título de exemplo, o sistema Greenstone¹¹, os sistemas do projeto OSS2lib¹² e o projeto CASPAR¹³. Porém, sistemas de software capazes de automatizar de maneira unificada informações de acervos históricos das mais diversas naturezas ainda não são encontrados. Nessa perspectiva, a norma internacional ISO 12127 - *Information and Documentation: A Reference Ontology for the Interchange of Cultural Heritage Information* (International Organization for Standardization - ISO, 2006) oferece uma base conceitual, no formato de uma ontologia, que possibilita intercambiar informações de acervos de naturezas distintas. Com base nessa norma, existem iniciativas de intercâmbio de informações, a exemplo do trabalho de Boeuf (2006), que apresenta um sistema para a troca de informações de acervos museológicos e bibliográficos. Contudo, vale ressaltar que a norma tem foco na necessidade de troca de informações e não na descrição completa de objetos pertencentes a qualquer tipo de acervo histórico. Além disso, ela não aborda questões relacionadas à implementação ou ao armazenamento das informações.

Na mesma perspectiva, um arcabouço foi proposto que define metadados para a descrição de bens patrimoniais, e busca assegurar que os dados representados sejam fiéis ao do bem patrimonial original (Beaudoin, 2012b). Esse arcabouço foi utilizado em um estudo de caso visando descrever dois bens patrimoniais, no caso uma pintura e uma ponte do século XIX. Segundo a autora, esse arcabouço deve ainda ser avaliado para outros tipos de bens patrimoniais. Além disso, conforme apontado por Beaudoin (2012b), o estabelecimento de tais metadados não é uma atividade trivial.

3 PADRÃO DE DESCRIÇÃO DA INFORMAÇÃO

O PDI é definido como um conjunto de atributos capaz de descrever completamente qualquer tipo de bem patrimonial, possibilitando a representação integrada de informações relativas a esses bens (CRISTIANINI et al., 2004). O

¹¹ <http://www.greenstone.org/>

¹² <http://www.oss4lib.org/projects>

¹³ <http://www.casparpreserves.eu/>

estabelecimento desse padrão foi resultado de extensas pesquisas de especialistas em Biblioteconomia, Museologia, Arquivologia, Ciência da Informação e Arquitetura, membros integrantes do Projeto Memória Virtual de São Carlos. Como discutido na Seção 2, nenhuma das normas investigadas atende de forma consensual às particularidades de todos os tipos de bens patrimoniais. Assim, cada uma das normas foi analisada, e foram identificados os elementos necessários para a descrição de cada tipo de bem. Em resumo, normas para a descrição de objetos bibliográficos e objetos de arquivos foram utilizadas como base para o PDI. Constatada a inexistência de normas para descrição de bens arquitetônicos, museológicos e naturais, recorreu-se a uma adaptação de roteiros de inventários, fichas cadastrais de diversas instituições e informações obtidas junto aos especialistas das áreas. Dentre as fontes que serviram de subsídios para a descrição desses bens incluem-se: fichas cadastrais dos bens do Museu de São Carlos, roteiro de tombamento do Conselho de Defesa do Patrimônio Histórico, Arqueológico, Artístico e Turístico (Conselho de Defesa do Patrimônio Histórico, Arqueológico, Artístico e Turístico - CONDEPHAAT, 1998), ficha de inventário dos bens arquitetônicos da Fundação Pró-Memória de São Carlos, roteiro para descrição de áreas verdes (FERNANDES; FURVINAL, 1997) e o Centro de Referência em Informação Ambiental (CRIA)¹⁴. A seguir, são listados os dez grupos de atributos que compõem o PDI:

- Grupo 1: Instituição - identificação da instituição e da unidade detentora dos acervos - compreende os atributos que representam os diversos dados de identificação da instituição que detém o bem patrimonial a ser descrito;
- Grupo 2: Bem patrimonial - identificação do bem patrimonial e nível de descrição - reúne os atributos que indicam o tipo do bem patrimonial que está sendo descrito, assim como a localização física do bem;
- Grupo 3: Título - compreende os atributos relacionados à variedade de tipos de títulos existentes para o bem patrimonial descrito;
- Grupo 4: Autoria - agrega os atributos relacionados à variedade de tipos de responsabilidades atribuídas à criação do bem patrimonial descrito;
- Grupo 5: Produção - contém os atributos que indicam o contexto no qual o bem patrimonial descrito foi produzido;
- Grupo 6: Descrição de conteúdo - compreende os atributos que representam a temática do bem patrimonial descrito e observações específicas;
- Grupo 7: Histórico e procedência - reúne os atributos que indicam a trajetória "de vida" do bem patrimonial descrito, bem como sua forma de aquisição;
- Grupo 8: Descrição física - agrega os atributos relacionados aos detalhes físicos e específicos de suporte do bem patrimonial descrito e também suas condições atuais de conservação e apresentação;

¹⁴ <http://www.cria.org.br/>

- Grupo 9: Condições de uso - contém os atributos que instruem sobre a possibilidade de acesso, reprodução, uso e proteção do bem patrimonial descrito; e
- Grupo 10: Responsáveis - Identificação dos responsáveis pelo produto informacional e compreende os atributos relacionados aos responsáveis pelo preenchimento dos dados sobre o bem patrimonial descrito, assim como sobre o contexto da pesquisa realizada e necessária para o correto preenchimento dos dados.

Esse padrão estabelece também “como” os dados serão armazenados na base de dados de um sistema de software, adotando, na medida do possível, normas internacionais de descrição de conteúdo, tais como as estabelecidas pelo AACR2 e pelo ISAD(G). Por exemplo, o atributo *autor* pertencente ao Grupo 4 (Autoria) deve ser escrito na forma direta, por extenso, de acordo com o documento ou a fonte de pesquisa (ex.: “Martinho Carlos de Arruda Botelho”). A título de ilustrar o PDI, na FIGURA 1 é apresentada parte do documento que descreve esse padrão.

Vale ressaltar que a definição de um conjunto de atributos comum para conjuntos distintos de informações não é uma tarefa trivial, uma vez que deve contemplar as especificidades desses vários conjuntos. Além disso, as normas internacionais foram instrumentos essenciais para a definição desse conjunto, bem como para buscar garantir a qualidade das informações armazenadas e disponibilizadas, possibilitando a integração com outros sistemas de informação.

PDI - Padrão de Descrição da Informação

Grupo 1: Instituição

...

Grupo 2: Bem Patrimonial

Tipo do Bem Patrimonial: indicar o item que está sendo inventariado.

Registro: número atribuído pela instituição ao bem patrimonial.

...

Grupo 3: Título

Título: título ou nome da obra, documento, bem arquitetônico, objeto museológico. Para os bens naturais, refere-se à equivalência para nome popular mais utilizado. Pode estar previamente estabelecido para o bem patrimonial ou será construído pelo pesquisador/catalogador ou ainda complementado de acordo com o plano de organização do acervo da instituição. Ex.: Título já estabelecido: “Revista Moderna”. Título construído: “Xícara do Conde do Pinhal”. Título complementado: “Revista Moderna fascículo 2”.

Título Original: obra ou documento traduzido, refere-se ao título na língua original; para os bens naturais, refere-se ao nome científico.

Título Alternativo: outro nome da obra ou documento que geralmente aparece acompanhado da conjunção “ou” após vírgula; para os bens naturais, refere-se à equivalência para outros nomes também populares ou regionais.

Acervo: nome da coleção a que pertence o item inventariado conforme organização da instituição.

Localização Física: conjunto de informações que descreve a localização física do bem patrimonial no acervo.

...

Grupo 4: Autoria

Autor: identificação da autoria deve ser escrita na forma direta, por extenso, de acordo com o documento ou a fonte de pesquisa. Ex.: Martinho Carlos de Arruda Botelho.

Função: função que a pessoa exerce em relação ao bem patrimonial. Ex.: escritor, proprietário, escultor, etc.

Tipo de Autoria: tipo de autoria:

- Autor: pessoa ou entidade autora ou criadora da obra, ou documento, ou projeto de edificação (por exemplo: arquiteto, engenheiro, projetista).

- Co-autor (autores com responsabilidades compartilhadas): pessoas ou entidades que participaram ou tiveram responsabilidades compartilhadas na produção ou geração de uma obra, documento, edificação, identificação de espécies. Exemplos: editor, diretor, tradutor, ilustrador, colaborador, coordenador, supervisor, revisor, compilador, organizador, adaptador, arranjador, relator, comentador, e intérprete (para som).

...

FIGURA 1: Parte do Padrão de Descrição de Informação

3.1 Modelagem do Padrão de Descrição da Informação

A partir do PDI, foram levantadas as características pertinentes a esse conjunto, bem como o relacionamento semântico entre os grupos de atributos, para elaborar um modelo capaz de expressar uma estrutura adequada da base de dados do sistema. Esse modelo foi construído utilizando-se o Modelo Entidade-Relacionamento (MER) (CHEN, 1976), conforme mostrado no Diagrama Entidade-Relacionamento (DER) da FIGURA 2. Para uma melhor visualização, são mostrados no DER apenas os elementos considerados como as entidades principais do PDI (ex.: BEM_PATRIMONIAL), os atributos que identificam essas entidades (ex: numero REGISTRO em BEM_PATRIMONIAL), os relacionamentos entre as entidades (ex: “detem” entre BEM_PATRIMONIAL e INSTITUICAO) e alguns atributos de entidades e relacionamentos. Além disso, a modelagem previu os vários tipos de usuários que o sistema iria dar suporte, bem como o seu comportamento perante os demais atributos da base. Buscou-se maximizar a independência entre os tipos de usuários do sistema (entidade USUARIO), o bem patrimonial descrito (entidade BEM_PATRIMONIAL) e a instituição detentora do acervo histórico (entidade INSTITUICAO), com o objetivo de flexibilizar o funcionamento do sistema.

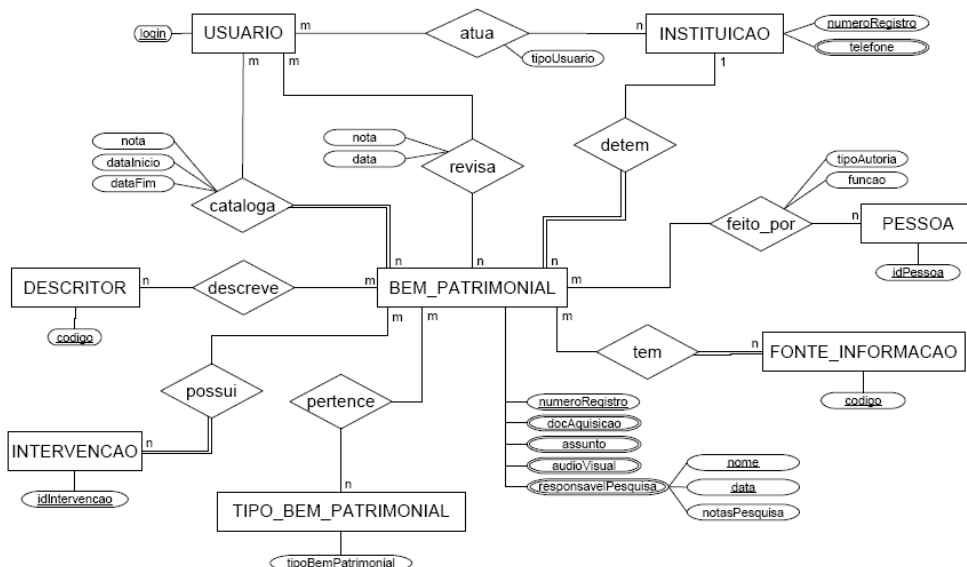


FIGURA 2 - Modelo da Base de Dados com suporte ao PDI

3.2 Validação do Padrão de Descrição da Informação

Sendo o PDI, bem como a estrutura da base de dados, elementos fundamentais para o desenvolvimento do sistema Memória Virtual, foi inicialmente construído um protótipo em MS Access¹⁵ para a validação desses elementos. Esse protótipo foi utilizado pelos grupos de trabalho do projeto — grupo de documentos bibliográficos, de objetos museológicos, de bens arquitetônicos, de bens naturais e de documentos de arquivo — com o intuito de validar a sua adequação para os diversos tipos de bens patrimoniais. Além disso, questões de usabilidade foram analisadas no protótipo, tais como a melhor forma de disponibilizar a grande quantidade de campos que precisariam ser preenchidos para cada bem patrimonial. Na FIGURA 3 é ilustrada uma das interfaces desse protótipo. Em seguida, com base nesse protótipo, o subsistema de catalogação de bens patrimoniais do Memória Virtual foi desenvolvido.

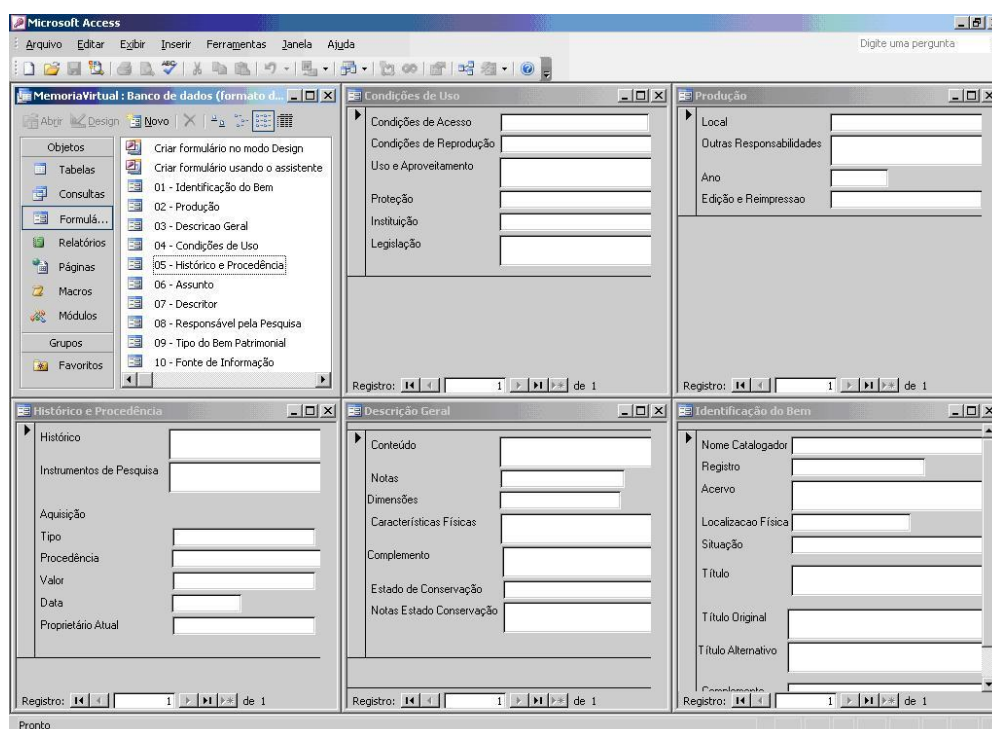


FIGURA 3- Protótipo de Validação do PDI

4 SISTEMA MEMÓRIA VIRTUAL

O Memória Virtual é um sistema disponibilizado na plataforma web que possibilita o armazenamento integrado e disponibilização de informações de acervos históricos, tendo como elemento principal o PDI. Considerando-se a relevância do movimento de software livre, esse sistema está licenciado sob a *General Public License*

¹⁵ <http://www.microsoft.com>

(GPL) (FSF, 1999), uma vez que essa licença é a mais adequada para garantir as liberdades principais de software livre: liberdade de executar, estudar, modificar e redistribuir o software, além de garantir que o software continue sendo livre. A seguir, são descritas as principais funcionalidades do sistema, bem como sua arquitetura, seu desenvolvimento, além das tecnologias que foram utilizadas para sua construção.

4.1 Funcionalidades do sistema

De modo geral, o Sistema Memória Virtual é composto de dois subsistemas: o subsistema de catalogação e o subsistema de acesso e navegação. O subsistema de catalogação possibilita o armazenamento de informações dos bens patrimoniais na base de dados do Memória Virtual por meio de uma interface web do usuário. As funcionalidades disponíveis nesse subsistema são: gerenciamento de usuários (administrador, gerente, revisor e catalogador), de instituições e de bens patrimoniais. Na FIGURA 4 é ilustrada a interface desse subsistema. Além das informações textuais referentes aos bens, a funcionalidade de *upload* de conteúdos multimídia está também disponível, possibilitando que imagens, vídeos e áudios sobre esses bens possam ser cadastrados, enriquecendo a base de dados do sistema. Além disso, em uma mesma instalação do sistema, informações de bens patrimoniais pertencentes a diversas instituições podem ser gerenciadas de forma integrada.

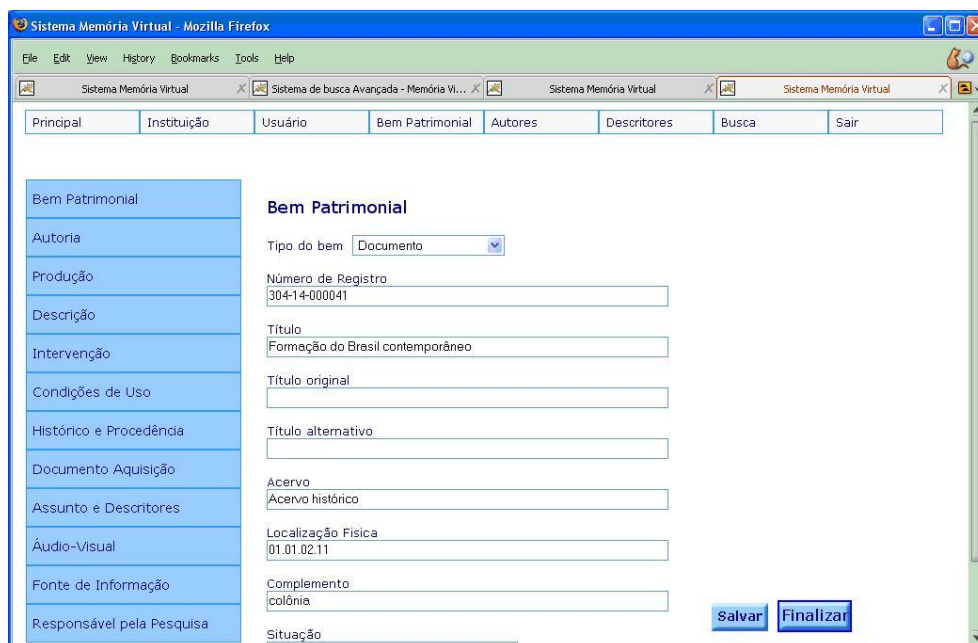
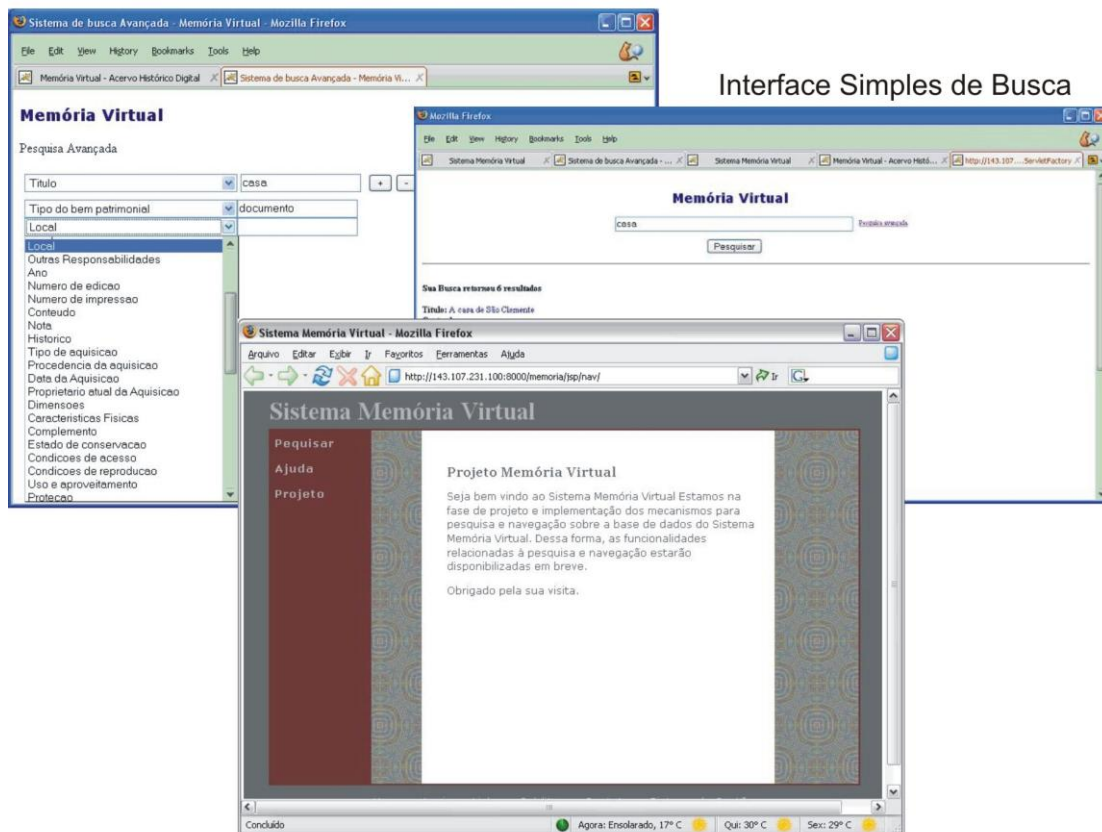


FIGURA 4 - Interface do Subsistema de Catalogação

O subsistema de acesso e navegação possibilita o acesso às informações dos bens patrimoniais contidos na base de dados, bem como a navegação entre essas informações. Ao contrário do subsistema de catalogação, esse subsistema apresenta interfaces distintas, ilustradas na FIGURA 5. Essa diversidade é necessária em função dos diferentes tipos de usuários para os quais o sistema se destina: de alunos do ensino fundamental a pesquisadores – visando o público em geral, uma interface no padrão

web também está disponível. Por meio dessas interfaces, é possível realizar uma busca simples (por um usuário cidadão em geral) ou uma busca avançada (por um usuário pesquisador). Os algoritmos de busca implementados consideram campos indexados (atributos do bem patrimonial que serão utilizados na busca pelo usuário cidadão e pelo usuário pesquisador). Foi também considerado o tratamento de *stopwords* (palavras que não devem ser consideradas na busca, tais como “a”, “de”, “com”, entre outros). Sabe-se que as informações armazenadas em bases de dados são elementos de grande valor em sistemas de software. Portanto, o compartilhamento dessas informações torna-se bastante relevante, principalmente quando se trata de sistemas no domínio de acervos históricos, propiciando a disseminação do conhecimento sobre esses acervos. O Memória Virtual disponibiliza então um módulo de exportação de dados, cujas interfaces são ilustradas na Figura 6. Esse módulo gera em XML o conteúdo referente aos dados armazenados na base. A XML foi adotada por mostrar-se uma das linguagens mais investigadas para a interoperabilidade de sistemas de software, principalmente em virtude de sua adaptabilidade a diversos domínios. Além disso, seguindo a filosofia de software livre e fazendo uma analogia de dados com código fonte de sistemas de software, uma licença de dados, no caso a *GNU Free Documentation License*¹⁶, é adotada para garantir que os dados exportados continuem sendo de acesso livre (NAKAGAWA, 2008).

Interface para Pesquisadores



Interface Padrão Web
FIGURA 5 - Interfaces do Subsistema de Acesso e Navegação

¹⁶ <http://www.gnu.org/licenses/fdl.html>

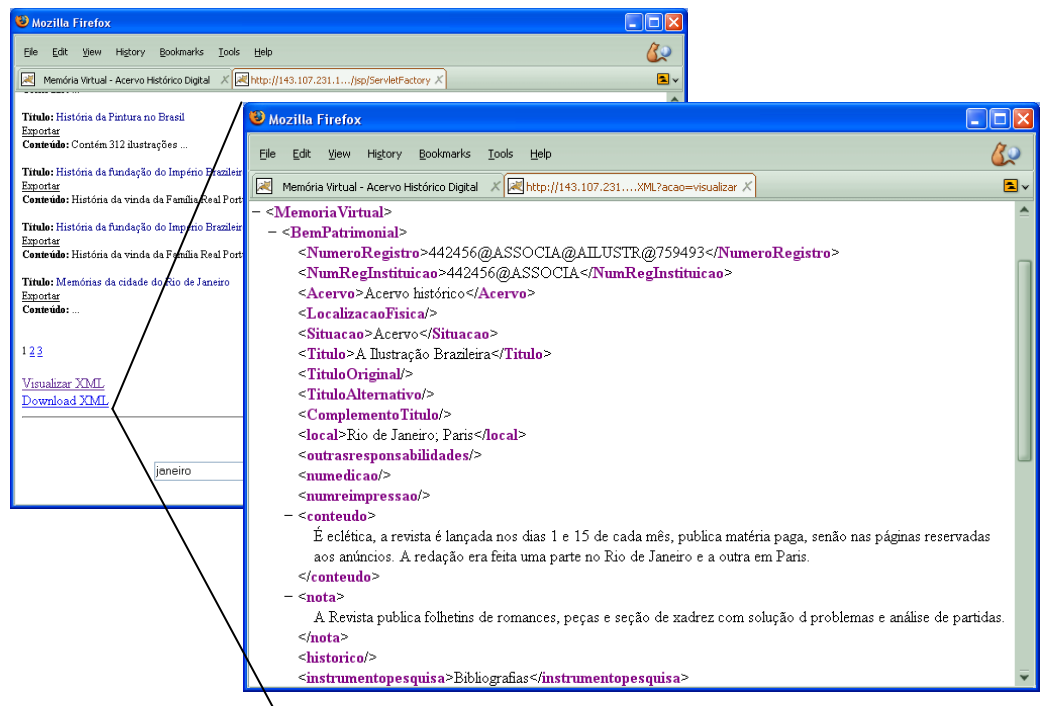


FIGURA 6 - Interface para Exportação de Dados

Em virtude das peculiaridades do Memória Virtual e visando a sua utilização adequada, bem como a qualidade das informações armazenadas na base de dados, *helps on-line* (ajudas) tornam-se indispensáveis e estão também disponíveis.

4.2 Projeto arquitetural do Memória Virtual

Uma das mais importantes preocupações, sendo o Memória Virtual um software livre, é sua capacidade de evolução e facilidade de adaptação para outras instituições que venham a utilizá-lo. Nessa linha, uma estruturação adequada da arquitetura do software desse sistema torna-se essencial. Para isso, foi conduzida uma investigação detalhada na área de arquiteturas de software, adotando-se como base arquiteturas mais recentes discutidas na literatura para o desenvolvimento de sistemas web (NAKAGAWA et al., 2005). Na FIGURA 7 é ilustrada a arquitetura do Memória Virtual, uma arquitetura baseada em três-camadas (do inglês, *three tier architecture*) (ECKERSON, 1995) composta da camada de apresentação, que contém a interface do usuário, da camada de aplicação, que contém a lógica de negócio, e da camada de persistência, cujo papel é gerenciar os objetos/elementos persistentes da aplicação. Essa arquitetura é também baseada no padrão arquitetural *Model-View-Controller* (MVC) (BUSCHMANN et al., 1996), no qual o Modelo (*Model*) contém as funcionalidades básicas e principais e gerencia os dados; a Visão (*View*) gerencia informações a serem mostradas para o usuário; e o Controlador (*Controller*) processa as entradas (eventos)

do usuário. Conforme ilustrado na FIGURA 7, a arquitetura do Memória Virtual é ainda composta de dois subsistemas — subsistema de catalogação e subsistema de acesso e navegação — apresentados na Seção 4.1.

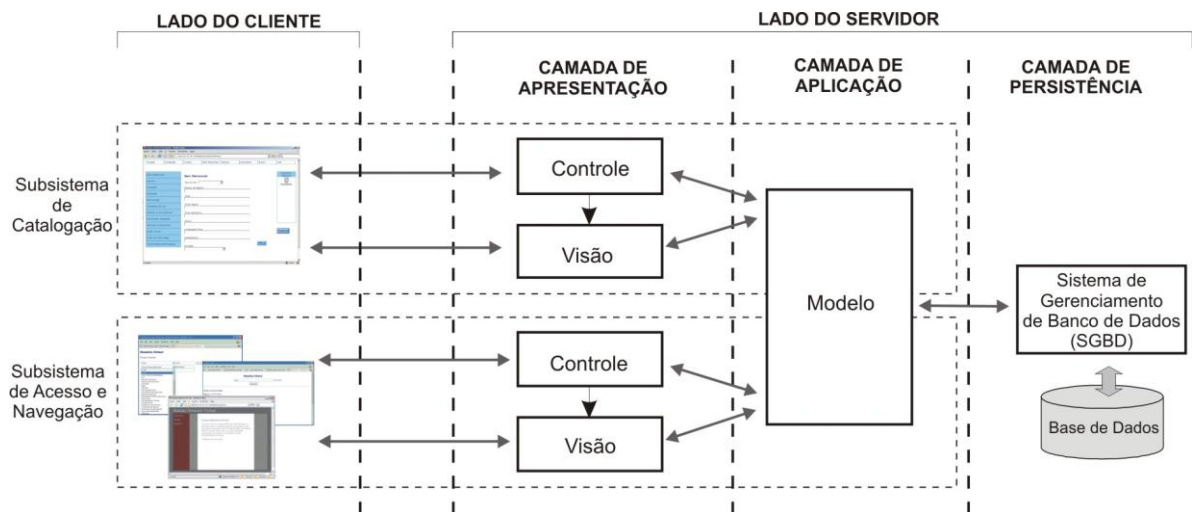


FIGURA 7 - Arquitetura do Sistema

4.3 Desenvolvimento do sistema

Para o desenvolvimento do Memória Virtual, foram adotadas técnicas e práticas que mais se adequavam a cada situação. Todo o desenvolvimento foi conduzido pelo grupo de desenvolvedores, que contou com a participação dos membros do projeto das diversas áreas relacionadas à Ciência da Informação, Museologia, Arquivologia, além de Ciências de Computação, que auxiliaram na definição dos requisitos do sistema, bem como no seu refinamento.

O desenvolvimento desse sistema, incluindo as fases iniciais de concepção até a fase de utilização, transcorreu em aproximadamente 30 meses, e diversos desenvolvedores atuaram no decorrer desse período. Ao todo, oito alunos de iniciação científica, além de outros oito alunos em projetos de graduação dos cursos de Bacharelado em Informática e Bacharelado em Ciências de Computação do ICMC/USP, envolveram-se no projeto e implementação do Memória Virtual.

Durante o período de desenvolvimento, diversas atividades foram conduzidas, sendo as principais delas: (i) estabelecimento dos requisitos; (ii) investigação de tecnologias livres para implementação e implantação do sistema; (iii) projeto arquitetural e da base de dados; (iv) implementação do sistema; e (v) validação do sistema.

A atividade de estabelecimento de requisitos envolveu um conjunto de exigências e particularidades que o próprio domínio exige. Por exemplo, o sistema exige padrões específicos de entrada, representação e armazenamento de dados. Portanto, um dos elementos mais importantes que delinearam os requisitos do sistema foi o PDI. A falta de sistemas semelhantes que servissem de “inspiração” e o

desconhecimento da equipe de desenvolvedores das características desse domínio de aplicação fizeram com que fosse necessária a utilização de uma abordagem baseada na coleta e gerência dos requisitos (NAKAGAWA; CRUZ; MALDONADO, 2006b). Para o levantamento, entendimento e refinamento dos requisitos, diversas estratégias foram experimentadas. Primeiramente, após a realização de *brainstorms* entre os membros do projeto, foi preparado um documento de requisitos baseado nas práticas recomendadas pela IEEE Std 830:1998 - *IEEE Recommended Practice for Software Requirements Specifications* (Institute of Electrical and Electronic Engineers - IEEE, 1998). Com base nesse documento, uma versão inicial do sistema foi desenvolvida e observou-se, por meio de frequentes reuniões com os clientes (membros do projeto), que a mudança nos requisitos era constante. Assim, ao invés de utilizar a abordagem de elaboração e consolidação de um documento de requisitos, os requisitos já eram implementados no sistema à medida que eram identificados. Ou seja, utilizou-se o próprio sistema em desenvolvimento para a identificação e refinamento de seus requisitos. Para essa atividade, as reuniões com os clientes foram de extrema importância para o refinamento do Memória Virtual.

A atividade de investigação de tecnologias livres para implementação e implantação do sistema teve como princípio a importância do ambiente de desenvolvimento de um software livre ser composto por tecnologias e ferramentas livres, o que facilitaria o posterior estabelecimento do mesmo ambiente para que o sistema possa sofrer alterações que se fizerem necessárias. Essa atividade resultou, portanto, em um conjunto de ferramentas para a construção do Memória Virtual. Frente à relevância dos resultados alcançados, um resumo dessas tecnologias é apresentado na Seção 4.4.

As atividades de projeto arquitetural e modelagem da base de dados resultaram na elaboração de modelos do sistema, tais como os já apresentados nas Figuras 2 e 7, que serviram de guia para a implementação. O sistema foi implementado a partir desses modelos, utilizando-se as tecnologias livres investigadas. Buscando ainda contribuir para a disseminação do Memória Virtual em nível internacional, na implementação do sistema foram aplicados os processos de internacionalização (do inglês, *internationalization* ou *i18n*) e localização (*localization* ou *L10n*) de software. A internacionalização refere-se a um processo em que se busca criar um software com características tais que permitam que ele se adapte mais facilmente a determinadas características regionais, legais, culturais e técnicas. Esse processo pode ser automático (de acordo com parâmetros recebidos, tais como fuso horário, formatos de números, sistemas métricos e moedas) ou manual (como a tradução da interface e adaptações específicas demandadas pelo destinatário do software), sendo que esse último é também chamado de localização de software. A interface do usuário do Memória Virtual tem sido então internacionalizada e localizada, sendo que a localização para a língua inglesa já foi realizada na estrutura da base de dados do sistema, considerando-se a universalidade do inglês, se comparado à língua portuguesa (SILVA, 2007).

Outra atividade foi o treinamento dos usuários visando à utilização adequada do sistema. A atividade de validação do sistema junto aos usuários foi conduzida objetivando refinar o Memória Virtual com a melhoria das funcionalidades já implementadas e identificação de novas funcionalidades.

4.4 Tecnologias livres para implementação e implantação do sistema

Nesta seção é apresentado um resumo de tecnologias que vêm sendo adotadas para o desenvolvimento de sistemas web e que foram utilizadas na construção do Memória Virtual. Adotou-se o PostgreSQL¹⁷ como Sistema de Gerenciamento de Banco de Dados (SGBD), considerando diversos critérios: sua licença, eficiência, segurança, capacidade de armazenamento e compatibilidade com outros softwares. Para a implementação das funcionalidades do sistema, foi escolhida a linguagem de programação *Java*¹⁸ e o Eclipse¹⁹ como ambiente de desenvolvimento. A escolha foi determinada basicamente por ser uma tecnologia relativamente recente, além de possibilitar a aplicação do paradigma Orientado a Objeto (OO). Além disso, para sincronizar o código fonte entre os desenvolvedores, registrar as modificações, suas datas e os responsáveis, adotou-se o SVN (Subversion)²⁰ como ferramenta de controle de versões.

Objetivando realizar um desenvolvimento mais padronizado e reutilizar soluções mais consolidadas em nível de projeto, investigou-se e utilizou-se um conjunto de padrões de projeto (do inglês, *design pattern*) (GAMMA et al., 1995), dentre eles o *Front Controller* e o *Data Access Object* (DAO). Além disso, visando ao aumento da produtividade no desenvolvimento, *frameworks* de aplicação foram investigados e utilizados, dentre eles o Struts²¹, que facilita a criação de aplicações web que utilizam a arquitetura MVC, e o Hibernate²², que trata do serviço de persistência e consulta objeto-relacional. As plataformas de software do ambiente de desenvolvimento e implantação são também livres: Linux²³ e Apache Tomcat²⁴. Em particular, o Apache Tomcat é um *web container* que implementa um ambiente de execução para os componentes especificados pela plataforma Java, incluindo segurança, concorrência, gerenciamento do ciclo de vida e transação.

É importante ressaltar que a seleção das tecnologias foi resultado de uma investigação rigorosa e justificada, estando todas elas alinhadas com as tecnologias mais recentes adotadas no desenvolvimento de sistemas web, com uma vasta comunidade de

¹⁷ <http://www.postgresql.org/>

¹⁸ <http://www.oracle.com/technetwork/java/index.html>

¹⁹ <http://www.eclipse.org/>

²⁰ <http://subversion.tigris.org/>

²¹ <http://struts.apache.org/>

²² <http://www.hibernate.org/>

²³ <http://www.linux.org/>

²⁴ <http://tomcat.apache.org/>

usuários e experiência consagrada na construção de uma diversidade de sistemas dos mais variados domínios de aplicação.

5 UTILIZAÇÃO DO SISTEMA

O levantamento de informações dos bens patrimoniais a serem catalogados no sistema Memória Virtual não é uma tarefa trivial, sendo um trabalho de pesquisa minucioso e demorado. Essa pesquisa deve ser ainda mais minuciosa se o objetivo for levantar pelos menos algumas das informações que tornem aquele bem patrimonial interessante quando disponibilizado para posterior consulta. Como mencionado na Seção 3, observa-se que cada bem patrimonial possui uma quantidade considerável de informações que devem ser identificadas, a saber, informações sobre autoria, descrição, produção, intervenção, condições de uso, histórico e procedência, assunto e descritores, documento de aquisição e áudio-visual. Para auxiliar essa tarefa, foi elaborado um manual de preenchimento que possibilita sistematizar o processo de coleta e descrição de informações de um bem patrimonial e, em seguida, a catalogação de tais informações no sistema.

Como estudos de caso, foram estabelecidos quatro projetos piloto que puderam contemplar vários tipos de bens patrimoniais que o sistema Memória Virtual propõe-se a gerenciar. Os acervos da Fazenda Pinhal²⁵, do Museu de São Carlos, da Fundação Pró-Memória de São Carlos e do Museu de Computação do ICMC/USP foram selecionados para serem catalogados no Memória Virtual.

A atual biblioteca da Fazenda Pinhal foi formada na década de 70, a partir da reunião dos livros que pertenceram à família do Conde do Pinhal, ao longo de muitas décadas. Da mesma forma, os documentos, tais como as fotografias, recortes de jornais, correspondências, entre outros (atualmente documentos do arquivo) encontravam-se em diferentes locais e móveis da casa grande. O acervo bibliográfico histórico, constituído por obras do começo do século XIX até a década de 70, é eclético, traduzindo os interesses variados dos membros da família ao longo dos anos. O conjunto de materiais bibliográficos da Fazenda Pinhal foi devidamente analisado (obra por obra) e dividido em três grandes tipos de acervos, de modo a facilitar a sua organização física e o seu uso, sem, entretanto, deixar de contar a sua história ou a sua procedência. Foram devidamente registrados no sistema Memória Virtual aproximadamente 1.500 bens patrimoniais, constituindo uma base de dados de tamanho razoável e com riqueza de conhecimento sobre os acervos da Fazenda Pinhal. O Museu de São Carlos possui quatro grandes coleções: histórica, fotográfica, etnográfica/artigos indígenas e artes visuais. Foram pesquisados e catalogados no Memória Virtual aproximadamente 2.500 bens patrimoniais. Já a Fundação Pró-Memória de São Carlos foi criada em 1993, com a finalidade de preservar e difundir o patrimônio histórico e cultural do Município de São

²⁵ <http://www.casadopinhal.com.br/>

Carlos. Entre as atividades que desenvolve estão as de reunir, conservar e disponibilizar o conjunto de seus documentos originários dos Poderes Executivo, Legislativo e Judiciário, além de outros advindos da administração pública direta e indireta, como também de fundos e coleções particulares, de reconhecido valor histórico e cultural. Nessa instituição foi realizado principalmente o tratamento da documentação relacionada aos bens arquitetônicos do município, que serão posteriormente catalogados no Memória Virtual. O Museu de Computação do ICMC-USP é constituído do acervo e da filosofia do “Museu de Instrumentos de Cálculo Numérico”, idealizado e criado em 1978 pelo Prof. Odelar Leite Linhares. Atualmente, o museu possui um acervo de mais de duas centenas de peças, entre máquinas de calcular, réguas de cálculo e computadores antigos. Foram catalogadas no sistema Memória Virtual 279 peças do museu, juntamente com suas imagens.

Esses estudos de caso têm demonstrado a capacidade do Memória Virtual e, conseqüentemente, do PDI, em tratar adequadamente bens patrimoniais de diversas naturezas. É relevante destacar ainda que não ocorreram situações nas quais a informação estava disponível e o sistema não possibilitou seu registro, por exemplo, pela falta de um campo adequado para aquela informação.

6 CONCLUSÃO

A contribuição principal deste trabalho é a apresentação dos principais resultados de um projeto de pesquisa multidisciplinar que envolveu pesquisadores, especialistas, profissionais e alunos de diversas áreas do conhecimento. Um resultado bastante relevante, inclusive em nível internacional, foi um padrão — o PDI — que possibilita a descrição completa e integrada de informações de acervos históricos de diversas naturezas. Além disso, outra contribuição deste trabalho é a descrição da experiência de desenvolvimento de um sistema web para o domínio de acervos históricos — o Sistema Memória Virtual — que tem como principal característica a automatização do PDI. Para tanto, foram apresentadas a arquitetura adotada, bem como as tecnologias livres utilizadas no desenvolvimento desse sistema. Quanto à catalogação de bens patrimoniais no sistema Memória Virtual, atualmente já estão cadastrados aproximadamente 4.300 bens dos mais variados tipos, o que mostra a versatilidade da base de dados do sistema. Além dos resultados tangíveis alcançados, vale ressaltar que ocorreu uma constante e enriquecedora troca de experiências entre as diversas áreas de pesquisa envolvidas, o que vem salientar a relevância da realização de projetos multidisciplinares. Do ponto de vista acadêmico, esse projeto foi relevante na medida em que proporcionou a formação de um grupo de estudos multidisciplinar, envolvendo professores, alunos e profissionais, sobre temas aglutinadores de diferentes áreas do conhecimento e que ensejam amplas possibilidades de pesquisa. Apesar dos muitos resultados positivos alcançados, deve-se ressaltar que o desenvolvimento de um sistema com as características do Memória Virtual e utilizando-se das tecnologias adotadas não

é uma tarefa trivial. Dessa forma, sendo desenvolvido como um dos resultados de um projeto de pesquisa, o Memória Virtual é ainda um sistema protótipo e que deve ser trabalhado para alcançar o status de um sistema comercial.

No contexto de um outro projeto de políticas públicas — Projeto Patrimônio Cultural Rural Paulista (PATRIMÔNIO, 2008) — tanto o Sistema Memória Virtual quanto o PDI foram evoluídos de modo a também abarcar informações dos acervos de fazendas históricas (COSTA, 2010; NAKAGAWA, 2010). Atualmente, esse sistema encontra-se em sua segunda versão e está sendo utilizado no projeto de pesquisa intitulado Critérios e Metodologias para Realização de Inventário do Patrimônio Cultural Rural Paulista (CRITÉRIO, 2012).

Como perspectivas de trabalhos futuros, pretende-se conduzir um estudo comparativo do PDI com as abordagens para a representação de bens patrimoniais, tais como as propostas por Beaudoin (2012b) e por Getty (2013). Isso possibilitará tanto evoluir o PDI, caso necessário, quanto coletar evidências da capacidade do PDI em contemplar um conjunto amplo de atributos que representem qualquer tipo de bem patrimonial, conforme sua proposta inicial. Além disso, como consequência da filosofia adotada, ou seja, do fato do Memória Virtual ser um sistema de software livre, podendo, portanto, ser adotado de forma gratuita, espera-se que esse sistema possa trazer benefícios para as instituições que queriam adotá-la e evoluí-la. Espera-se, também, benefícios para toda a comunidade, uma vez que o sistema poderá viabilizar a disponibilização integrada de informações de acervos históricos dos mais diversos tipos, contribuindo inclusive para o registro histórico dos diversos acervos nacionais.

Agradecimentos:

A todos os membros do Projeto Memória Virtual de São Carlos e em especial à Prefeitura Municipal de São Carlos, à Associação Pró Casa do Pinhal, ao Museu de São Carlos e à Fundação Pró-Memória de São Carlos. Agradecemos ainda aos pesquisadores, professores e alunos do Instituto de Ciências Matemáticas e de Computação e do Departamento de Arquitetura e Urbanismo da Universidade de São Paulo, e do Departamento de Ciência da Informação da Universidade Federal de São Carlos. Agradecemos também à FAPESP, Capes e CNPq pelo apoio financeiro.

REFERÊNCIAS

ANGLO-AMERICAN CATALOGUING RULES. Anglo-american cataloguing rules. 2. ed. Disponível em: <<http://www.aacr2.org/>>. Acesso em: 18.02.2014.

BEAUDOIN, J. E. Context and its role in the digital preservation of cultural objects. **D-Lib Magazine**, v. 18, n. 11/12, p. 1-15, Novembro/Dezembro 2012a.

BEAUDOIN, J. E. A framework for contextual metadata used in the digital preservation of cultural objects. **D-Lib Magazine**, v. 18, n. 11/12, p. 1-15, Novembro/Dezembro 2012b.

BOEUF, P. L. Using an ontology-driven system to integrate museum information and library information. In: SYMPOSIUM ON DIGITAL SEMANTIC CONTENT ACROSS CULTURES, 1., 2006, Paris. **Proceedings** ... Paris: the Louvre, 2010. p. 1-4.

BUSCHMANN, F., et al. **Pattern-oriented software architecture: a system of patterns**. New York: John Wiley & Sons, 1996. 457 p.

CHEN, P. P. The entity-relationship model - toward a unified view of data. **ACM Transactions on Database Systems**, v. 1, n. 1, p. 9-36, 1976.

CONSELHO DE DEFESA DO PATRIMÔNIO HISTÓRICO, ARQUEOLÓGICO, ARTÍSTICO E TURÍSTICO. **Patrimônio cultural paulista, CONDEPHAAT: bens tombados 1968–1998**. São Paulo : Imprensa Oficial do Estado, 1998.

COSTA, L. S. F., et al. Hacienda históricas en el interior paulista de los siglos XVIII y XIX: una propuesta teórico y el metodológico de inventario del patrimonio. In: CONGRESO INTERNACIONAL DE REHABILITACIÓN DEL PATRIMONIO ARQUITECTÓNICO Y EDIFICACIÓN, 10., 2010, Santiago, Chile. **Anais...** Chile: CICOP, 2010. p. 1-6.

CRISTIANINI, G. M. S., et al. Conteúdos digitais e padrões de registros: desafios para a democratização de acervos especiais. In: SEMINÁRIO NACIONAL DE BIBLIOTECAS UNIVERSITÁRIAS, 13., 2004, Natal. **Anais...** Natal: SNBU, 2004. p. 1-14.

CRITÉRIOS e metodologias para realização de inventário do Patrimônio Cultural Rural Paulista, Universidade Federal de São Carlos (UFSCar), São Carlos, SP, 2012-2014. (Projeto aprovado pela FAPESP e coordenado pelo Profa. Dra. Luzia Sigoli Fernandes Costa)

ECKERSON, W. W. Three tier client/server architecture: achieving scalability, performance, and efficiency in client server applications. **Open Information Systems Journal**, v. 10, n. 1, p. 1-20, jan. 1995.

FERNANDES, L. S.; FURVINAL, A. C. Informação para educação ambiental: proposta para a criação de um centro de referência. In: SEMINÁRIO DE CIÊNCIA E DESENVOLVIMENTO SUSTENTÁVEL, 1., 1997, São Paulo. **Anais...** São Paulo: IEA/USP, 1997.

FREE SOFTWARE FOUNDATION. **The Free Software definition**. 2013. Disponível em: < <http://www.gnu.org/philosophy/free-sw.html> > Acesso em: 18.02.2014.

FREE SOFTWARE FOUNDATION. In: OPEN SOURCES: VOICES FROM THE OPEN SOURCE REVOLUTION. **GNU General Public Licence**. Beijing, 1999, v.1, p. 258–262.

GAMMA, E., et al. **Design patterns – elements of reusable object-oriented software**. Reading: Addison-Wesley, 1995. 395 p.

GETTY, P. TRUST & COLLEGE OF ART ASSOCIATION. **Categories for the description of works of art**. Disponível em <http://www.getty.edu/research/publications/electronic_publications/cdwa/>. Acesso em: 18.02.2014.

INTERNATIONAL COUNCIL ON ARCHIVES. **ISAD(G): general international standard archival description**, 2. ed. Disponível em <<http://www.ica.org/10207/standards/isadg-general-international-standard-archival-description-second-edition.html>>. Acesso em: 18.02.2014.

INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS. **IEEE 830 – IEEE recommended practice for software requirements specifications**. 1998.
INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 21127:2006. Information and documentation – a reference ontology for the interchange of cultural heritage information**. Geneva, 2006.

LIBRARY OF CONGRESS. **Marc standards**. Disponível em <<http://www.loc.gov/marc/>>. Acesso em: 18.02.2014.

MEMÓRIA Virtual de São Carlos. Instituto de Ciências Matemáticas e de Computação - ICMC/USP, São Carlos, SP, 2004-2008. (Projeto aprovado no programa de políticas públicas da FAPESP e coordenado pelo Prof. Dr. José Carlos Maldonado).

NAKAGAWA, E. Y., et al. Disseminação de informações via dados livres. In: **WORKSHOP SOBRE SOFTWARE LIVRE, FÓRUM INTERNACIONAL SOFTWARE LIVRE**, 9., 2008, Porto Alegre. **Anais...** Porto Alegre: SBC, 2008. p. 1-6.

NAKAGAWA, E. Y., et al. Memória virtual: estabelecimento da arquitetura e requisitos do sistema. In: **WORKSHOP SOBRE SOFTWARE LIVRE, FÓRUM INTERNACIONAL SOFTWARE LIVRE**, 6., 2005, Porto Alegre. **Anais...** Porto Alegre: Armazem Digital, 2005. p. 1-6.

NAKAGAWA, E. Y.; CRISTIANINI, G.M.S.; MORAES, J.S. Democratização de acervos especiais: o Modelo do Projeto “Memória Virtual de São Carlos”. In: **SEMINÁRIO NACIONAL DE BIBLIOTECAS UNIVERSITÁRIAS**, 14., 2006, Salvador. **Anais...** Porto Alegre: SBC, 2006a. p. 1-10.

NAKAGAWA, E. Y.; CRUZ, N. F.; MALDONADO, J. C. Relevância dos requisitos no desenvolvimento de software livre. In: **WORKSHOP SOBRE SOFTWARE LIVRE, FÓRUM INTERNACIONAL SOFTWARE LIVRE**, 7., 2006, Porto Alegre. **Anais...** Porto Alegre: SBC, 2006b. p. 171-176.

NAKAGAWA, E. Y., et al. Fazendas históricas paulistas dos séculos XVIII e XIX: premissas teóricas e metodológicas para inventariar bens patrimoniais. **Revista Resgate**, Unicamp, v. 18, n. 20, p. 37-53, 2010.

PATRIMÔNIO Cultural Rural Paulista, Centro de Memória da UNICAMP, Campinas, SP. 2008-2011. (Projeto aprovado no programa de políticas públicas da FAPESP e coordenado pelo Prof. Dr. Marcos Tognon).

RAY, E. T. **Learning XML**. 2. ed. O'Reilly, 2003. 416 p.

SILVA, T. C. **Aplicação de localização de software na estrutura da base de dados do Sistema Memória Virtual**. 2007. 69 f. Monografia (Projeto de Graduação) - Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos.

Como citar este documento:

NAKAGAWA, Elisa Yumi et al. Um Sistema Livre para Automatização do Padrão de Descrição da Informação. **Rev. digit. bibliotecon. cienc. inf.**, SP, v.12, n.2, p.173-192, maio/ago. 2014. ISSN 1678-765X. Disponível em: www.sbu.unicamp.br/seer/ojs/index.php/rbci>. Acesso em: 30 maio 2014