

Ambiente para busca e visualização de documentos históricos na Web

Maurilio de Araujo Possi

**Mestrando em Ciência da Computação.
Universidade Federal de Viçosa**

Alcione de Paiva Oliveira

Professor Associado. Universidade Federal de Viçosa Departamento de Informática

Alexandra Moreira

Doutoranda em Lingüística Universidade Federal de Juiz de Fora

Fábio Mendes

Professor Associado. Universidade Federal de Viçosa Departamento de História

Jonas Marçal Queiroz

Professor Adjunto-Universidade Federal de Viçosa Departamento de História

Documentos históricos são ferramentas essenciais para historiadores. Em muitos casos, o acesso aos documentos pode ser difícil em função de inúmeros fatores, tais como a distância, segurança e fragilidade do documento. Uma forma de contornar esse problema é a digitalização e disponibilização do acesso em uma rede de comunicação, a Internet. Este artigo apresenta um ambiente para visualização de documentos históricos, desenvolvido dentro do projeto de Digitalização de manuscritos dos acervos dos Cartórios do 1º. e 2º. Ofícios do Arquivo Histórico da Casa Setecentista de Mariana - PHAN.

Palavras-chave: Digitalização; Documentos históricos; Busca na Web.

Web environment for search and display of historical documents

Historical documents are essential tools for historians and researchers. However, in many cases, access to documents can be difficult or impossible, depending on factors such as distance, security and fragility of the document. One way to circumvent this problem would be through scanning and allowing access through an open network of communication, such as the Internet. This article presents an environment for viewing of historical documents developed within the digitalization project of the manuscripts of the historical archive of the "Cartórios do 1º e 2º Ofícios Casa Setecentista de Mariana-IPHAN".

Keywords: Digitization; Historical Documents; Web Search.

Recebido em 14.07.2009 Aceito em 30.07.2011

1 Introdução

A digitalização tem sido vista como um caminho para complementar soluções relacionadas tanto com a preservação como ao acesso de documentos históricos, como sugerido por MOREIRA *et al.* (2007). É um caminho complementar, porque ainda não é possível substituir a técnica da microfilmagem; no entanto, a digitalização torna possível restringir o acesso ao documento original, liberando para consulta apenas o material digitalizado. Além disso, a possibilidade da transmissão das informações digitalizadas por meio de uma rede de computadores, como a Internet, permite um maior acesso e agilidade às solicitações de conteúdo pelo público. Esse último ponto é fundamental para a universalização do acesso aos documentos e ao incremento das pesquisas, no âmbito da história e da arquivística. No entanto, é importante que a disponibilização dos documentos atendam a alguns requisitos, de forma a atender adequadamente à sua finalidade. Existem vários requisitos de segurança, como desempenho, qualidade de imagem e metadados para a busca, que serão detalhados na próxima seção; eles devem ser atendidos por ambientes que tem como finalidade prover acesso remoto a documentos históricos.

Este artigo relata a experiência do projeto e implementação de um ambiente para disponibilização de documentos históricos, que usa como meio de comunicação a Internet. Esta atividade ocorreu dentro do projeto de Digitalização de manuscritos dos acervos dos Cartórios do 1º. e 2º.

Ofícios do Arquivo Histórico da Casa Setecentista de Mariana - IPHAN, coordenado pelo Departamento de Artes e Humanidades da Universidade Federal de Viçosa e financiado pela FAPEMIG, por meio do edital FAPEMIG No.009/2005 - "USO DA TECNOLOGIA DIGITAL NO RESGATE DA IDENTIDADE HISTÓRICO-CULTURAL DE MINAS GERAIS" (MOREIRA *et al.*, 2007; MENDES *et al.*, 2007). Espera-se que este relato ajude a nortear projetos semelhantes e que os critérios adotados possam auxiliar nas decisões difíceis, que surgem em projetos desta natureza.

A próxima seção descreve os requisitos em ambientes com essa finalidade. Em seguida, é apresentada a arquitetura e funcionalidades do ambiente desenvolvido. A seção 4 apresenta os resultados obtidos com o ambiente. Finalmente, a seção 5 apresenta as conclusões relativas ao projeto e implementação deste ambiente.

2 Requisitos para Busca e visualização de documentos históricos à distância

Os seguintes requisitos devem ser atendidos por ambientes que têm por finalidade a disponibilização de documentos históricos à distância:

desempenho: documentos históricos, especialmente manuscritos, fotografias e gravuras, são digitalizados e arquivados no formato de arquivos imagens, como o formato *Tagged Image File Format* (TIFF) ou *Joint Photographic Expert Group* (JPEG). Em função da necessidade de se realizar uma digitalização a mais próxima possível do documento original, o processo é realizado utilizando-se uma grande resolução (mínimo de 300 ppi) e profundidade de bits (24 bits). Portanto, esses arquivos ocupam uma grande quantidade espaço em meios de armazenamento. O tamanho típico de apenas uma página do projeto fica em torno de 15 *megabytes*. A transmissão de arquivos desse tamanho, via uma rede de computadores de longa distância é, para os padrões atuais de transmissão, inviável. Além disso, a digitalização nessa resolução e profundidade de bits é necessária apenas para armazenamento e realização de futuras cópias, sem a necessidade de se recorrer ao documento original. Para pesquisas comuns, imagens com resolução bem menores, atendem às necessidades dos usuários. Geralmente, o requisito do desempenho é mais prioritário e, para atendê-lo, é necessário reduzir a resolução das imagens, diminuindo seu tamanho para apenas alguns *megabytes*, para que possam ser transmitidas rapidamente pela Internet, mas, de tal forma, que seu conteúdo continue fiel ao original, não apresentando distorções, e possa, assim, serem vistas com facilidade. Além disso, é necessário prover meios de pré-visualização, como, por exemplo, a exibição de miniaturas (ou *thumbnails*), para que o usuário possa navegar rapidamente pelas imagens até que encontre uma imagem específica que seja do seu interesse, e aí sim, acesse o seu conteúdo em tamanho original;

a)segurança: ao disponibilizar acesso a documentos históricos em um meio público e tão irrestrito como a Internet, segurança é um grande

requisito e, talvez, o maior desafio. Existem diversos motivos pelos quais se torna necessária a segurança e o controle de acesso em sistemas como esse, dos quais se pode salientar a possibilidade de tais documentos estarem protegidos por leis de direitos autorais, que proíbem sua cópia por terceiros. Para atender o requisito de segurança, o sistema deve ser capaz de identificar os usuários do sistema e possuir algum meio para garantir a verificação da origem do documento;

b) atendimento aos padrões arquivísticos: o arquivamento de documentos de qualquer natureza deve seguir as normas estabelecidas pela organizações relacionadas com a arquivística, como o Conselho Nacional de Arquivos (CONARQ), de modo a estabelecer a proveniência, facilitar o acesso e o intercâmbio de documentos. A norma de arquivamento mais aceita atualmente é a *International Standard Archival Description* (ISAD-G) (INTERNATIONAL COUNCIL ON ARCHIVES - ICA, 2000). A ISAD(G) é uma norma estabelecida pelo Conselho Internacional de Arquivos, que visa descrever um conteúdo informacional, aplicando-se um sistema de níveis. Esse sistema de níveis permite uma descrição detalhada e racional, facilitando o entendimento e o acesso mútuo. Além disso, respeita os princípios da proveniência e da organicidade; e

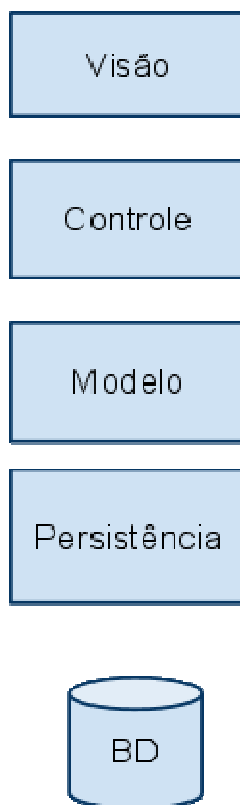
c) acessabilidade: o requisito da acessabilidade está relacionado à facilidade de uso e de busca do sistema. Para cumprir seu papel de ferramenta para universalização do acesso à informação, o ambiente de acesso aos documentos deve ser de fácil utilização e, também, permitir que o usuário possa realizar vários tipos de busca. Os documentos exibidos também devem ser apresentados de forma que facilite a visualização e a navegação.

3 Soluções adotadas para o atendimento dos requisitos

Os requisitos levantados serviram de base para orientar a definição da arquitetura do sistema, que recebeu a denominação de Sistema de Arquivamento e Pesquisa de Documentos Históricos (DigiDoc).

Foi projetada uma arquitetura multicamada, utilizando o padrão arquitetural Modelo, Visão e Controle (MVC), que preconiza a separação de responsabilidades em camadas de software. Cada camada de software usa os serviços da camada subjacente e oferece serviços para as camadas superiores. A divisão das responsabilidades em camadas de software torna o sistema mais fácil de ser modificado e mais escalável. O padrão MVC determina que o sistema possua pelo menos três camadas de responsabilidades: a camada da visão, responsável pela interface com o usuário; a camada de controle, responsável pela captura das requisições do usuário e por sua execução; e a camada do modelo, responsável pela manutenção do modelo de informação da aplicação. Em sistemas mais complexos, é necessário introduzir um número maior de camadas de softwares para reduzir a complexidade geral da aplicação. A FIG. 1 mostra as camadas usadas na aplicação.

Figura 1 - Camadas do sistema



Além das camadas mínimas do MVC, foi adicionada uma camada de persistência, responsável para fornecer os serviços de persistência e recuperação de informação no Banco de Dados. A arquitetura em camadas proporcionou uma maior facilidade para atender aos requisitos, uma vez que cada um deles podia ser tratado separadamente, em cada camada. A seguir, serão descritas as abordagens adotadas para atendimento dos requisitos listados na seção anterior.

No atendimento do requisito desempenho, as seguintes abordagens foram realizadas: adoção da arquitetura em camadas; redução da resolução das imagens dos documentos; e o uso de ferramenta dinâmica de zoom. A arquitetura em camadas favorece ao desempenho, por permitir alocar cada camada em uma máquina separada. Com isso, pode-se adquirir máquinas de acordo com a necessidade em atender à demanda de acesso. Isto ainda não foi necessário e, portanto, o sistema está sendo executado em um único servidor.

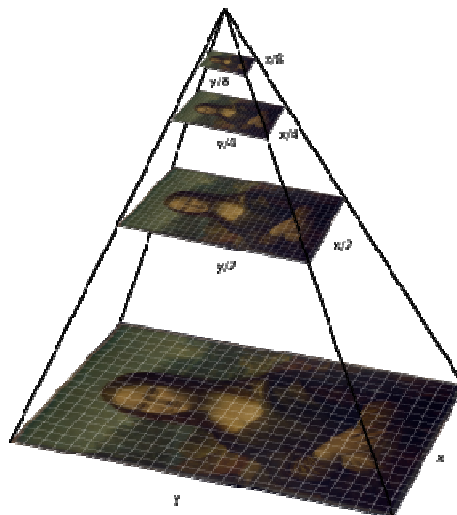
A imagem original foi reduzida para cerca de um *megabyte*, tanto para facilitar a transmissão quanto para evitar que a imagem seja utilizada como uma imagem fiel do documento original. Ainda assim, o tamanho de

um *megabyte* é muito grande para a maioria das conexões na Internet. Para solucionar esse problema, foi adotado o formato *pyramidal TIFF*.

O formato *Pyramidal TIFF* consiste em várias camadas de imagem (*layers*), cada uma utilizando uma resolução diferente (*multi-resolution*). Cada camada é separada em pequenos blocos (*tiled*). As camadas são ordenadas por sua resolução, resultando em uma forma semelhante a de uma pirâmide, daí o seu nome. Esse formato permite o acesso instantâneo a qualquer bloco individual de imagem, em qualquer resolução (a resolução pretendida pode variar de acordo com o nível de aproximação (zoom) que o usuário estiver usando), sem que seja necessário transmitir toda a imagem, minimizando a sobrecarga da rede e do servidor. Na medida em que o usuário vai navegando pela imagem, novos blocos vão sendo requisitados e a imagem vai se montando, parte a parte. Se o usuário voltar para uma região onde os blocos já foram recebidos, eles não serão carregados do servidor novamente, uma vez que já se encontram no *cache* do navegador. Ao final, caso o usuário tenha visto toda a extensão da imagem, terá recebido o arquivo todo, sem ter esperado pelo recebimento completo do mesmo.

O *Pyramidal TIFF* é uma extensão padrão do formato TIFF e é reconhecido pela maioria dos programas de edição gráfica, como o GIMP¹, VIPS², ImageMagick³, etc. Para converter as imagens originais nesse formato, foi utilizado o programa VIPS, um sistema livre de processamento de imagens, integrado ao sistema web, de modo que a conversão é feita automaticamente, sempre que se adiciona as imagens de um novo documento ao sistema.

Figura 2 - *Pyramidal TIFF*



Fonte: INTERNET IMAGING PROTOCOL – IIP (1997).

1 Disponível em: <www.gimp.org>. Acesso em: 12 mar. 2010.

2 Disponível em: <www.vips.ecs.soton.ac.uk>. Acesso em: 03 maio 2009.

3 Disponível em: <www.imagemagick.org>. Acesso em: 03 maio 2009.

As imagens no formato *Pyramidal TIFF* são enviadas para página do cliente via o protocolo *Internet Image Protocol (IIP, 1997)*. Para o envio das imagens via IIP, foi utilizado o programa servidor IIPImage, que comunica com a página por meio de requisiões dinâmicas em *Asynchronous Javascript And XML (AJAX)* (GARRETT, 2005).

No atendimento do requisito segurança, foi usada a estratégia de autorização acesso, por meio de login baseada em *e-mail* válido. Para isso, o usuário deve se cadastrar no site, fornecendo seus dados pessoais, filiação e um *e-mail* válido, que receberá uma mensagem eletrônica, contendo um *link* que deve ser acessado para a confirmação do cadastro. Esta estratégia evita que o usuário forneça um *e-mail* falso no cadastro e permite que o usuário seja localizado por meio do provedor do serviço de *e-mail*. Isso não evita que um usuário crie um email com dados falsos em um provedor de *e-mail*, mas garante que, pelo menos, um acesso foi realizado nesse provedor, deixando um algum registro. Uma verificação de identidade mais rígida chegou a ser cogitada, como o cadastramento via apresentação de documentos, mas essa abordagem foi descartada, em nome de uma maior facilidade de acesso aos documentos, o que é um dos principais objetivos da criação do site.

Outra restrição é o impedimento de baixar, na íntegra, as imagens dos documentos. Os documentos são exibidos em uma janela do *site*, por meio da tecnologia AJAX, que mostra apenas parte do documento, com uma marca d'água no canto inferior direito (FIG. 3).

Figura 3 - Apresentação dos documentos digitalizados



Essa forma de apresentação evita que o documento seja usado como uma imagem fiel do original e atende apenas aos propósitos de pesquisa.

Para o atendimento do requisito acessibilidade, é apresentada ao usuário diversas formas de acesso aos documentos. A forma básica de acesso aos documentos é por meio da navegação estrutura hierárquica (FIG. 4), que é organizada, segundo os níveis que refletem a organização interna do arquivo, ou seja: cartório (primeiro ou segundo ofício); tipo de documento (ação cível, crime, execução, inventário, justificação, notificação, sesmaria, testamentaria, alforrias–cartas, audiências, compra e venda de escravos, fianças de criminosos, hipotecas, notas, procurações, querelas, registro do testamento); ano; número do códice; e número do auto. Atualmente, somente os inventários do primeiro ofício no período de 1713 até 1770 estão disponíveis.

Figura 4- Árvore de navegação



O sistema permite, também, que o usuário realize buscas sobre os metadados dos documentos. A busca pode ser tanto na modalidade exata quanto na modalidade aproximada.

Na busca exata, os termos de busca fornecidos pelo usuário são passados para o SGBD, no caso o MySQL4, que se encarrega de fazer a busca nos metadados dos documentos cadastrados. Essa busca é chamada de exata, pois os termos cadastrados no banco têm que ser idênticos aos termos digitados para serem considerados nos resultados da busca. Por exemplo, os documentos que façam referência a uma pessoa chamada "Thiago Costa da Silva", serão encontrados se o usuário buscar por "Thiago", mas não serão se o mesmo usuário digitar "Tiago", pois as duas palavras diferem-se pela letra "h". Apesar disso, decidiu-se que a busca exata aceitaria subsequências de palavras e não só palavras inteiras, isto é, a busca pelo termo "arma" poderia ter como resultados quaisquer documentos que tenham palavras com a sequência "arma", como, por exemplo, "armando", "armador", "farmácia", etc. A busca também não é sensível ao caso maiúsculo ou minúsculo dos caracteres, ou seja, os documentos que façam referência a uma pessoa chamada "Antônio Carvalho da Cunha" poderão ser encontrados caso o usuário digite "Antônio", "antônio", "ANTÔNIO", etc. E, ainda, a busca exata vai considerar como resultado palavras que diferem-se por acentos ou cedilha. Em resumo, documentos que têm a palavra "Antônio" poderão ser encontrados pelo usuário que digite "antonio", "antônio", "Antonio", "Antônio", "Anton", "Ant", etc., na busca exata. A desvantagem de se usar a busca exata é que torna-se necessário saber com certa exatidão o termo que se está procurando. Além disso, termos com erros ortográficos ou com alterações na grafia não serão encontrados. A grande vantagem é a rapidez desse método, capaz de varrer todos os metadados em alguns milissegundos, dependendo da quantidade de documentos cadastrados.

Como os documentos podem ser de qualquer época e as línguas, em especial a portuguesa, vem passando por constantes transformações, surgiu um inconveniente: a grafia de muitos nomes foram alterados com o passar do tempo. Além disso, a transcrição e o cadastro dos documentos são passíveis de erros humanos, incluindo o de digitação, que podem modificar a grafia de algum termo. Apesar de toda flexibilidade que foi dada à busca exata, ela ainda não é capaz de encontrar documentos ignorando certas mudanças de grafia, seja por erro humano, seja por evolução da língua. Para solucionar este problema, propôs-se a implementação de uma busca aproximada.

Na busca aproximada, o sistema permite que o usuário a realize por termos aproximados, utilizando o algoritmo shift-and para casamento aproximado, proposto por Ziviani (2004). Neste algoritmo, a busca é modelada como um autômato finito não determinístico e permite que ocorra casamento de palavra, com até um certo número de erros ou distância da palavra original. Distância de edição k entre duas cadeias de caracteres x e y é o menor número de operações de inserção, edição ou remoção necessárias para converter x em y ou vice-versa. Na implementação usada, é permitido o casamento com até três erros e o tamanho mínimo da palavra é de seis caracteres.

4 Disponível em: <www.mysql.com>. Acesso em: 11 fev. 2009.

O mais importante para o usuário, entretanto, são os recursos flexíveis de visualização desenvolvidos para o DigiDoc, desenhados com vista às dificuldades da leitura paleográfica de manuscritos. A ferramenta de visualização permite o zoom da imagem com clique-duplo sobre a imagem ou através do uso do roller do mouse. A miniatura que aparece à direita da imagem permite o deslocamento da imagem e a seleção de parte do documento para leitura. A barra de rolagem à esquerda, permite ao consulente uma noção precisa do número total de páginas do auto, assim como a transição rápida para qualquer parte do mesmo. Os ícones na parte superior permitem modificar a cor de fundo e/ou visualizar o documento em negativo.

No atendimento do requisito "atendimento aos padrões arquivísticos", adotou-se um conjunto de metadados descritivos, selecionados de um subconjunto dos que são sugeridos na norma ISAD(G) (ICA, 2000). A norma ISAD (G) não é uma norma prescritiva, mas sim, diretiva e, por esse motivo, foram usadas como base para a escolha dos metadados adequados ao projeto.

Os metadados aplicados aos documentos arquivados na Casa Setecentista de Mariana – MG, são de dois tipos: os de acesso público e os de acesso restrito. Os metadados de acesso restrito, tais como procedência e condições de reprodução, são utilizados pelos funcionários ou técnicos, nas atividades de descrição do conteúdo da documentação, enquanto que os de acesso público, tais como, idioma, título, âmbito e conteúdo, não possuem restrição de acesso. Outra classificação para os metadados, proposta pela ISAD(G), é separação em áreas, que agrupam os elementos de descrição comuns.

Os metadados foram digitados em uma planilha, para, posteriormente, serem inseridos em um banco de dados. Está prevista para futuras versões do sistema a exportação dos dados no formato Encoded Archival Description (EAD), que é um padrão de marcação para descrição de documentos, apoiado pela Sociedade Americana de Arquivistas e pela Biblioteca do Congresso Americano.

5 Resultados

O release beta do sítio foi oficialmente lançado no dia 10 de outubro de 2008, na cidade de Mariana⁵ (FIG. 5). Na oportunidade, estiveram presentes diversas autoridades da Universidade Federal de Viçosa e do IPHAN, bem como convidados.

As imagens, porém, não foram ainda disponibilizadas na sua totalidade. As imagens já digitalizadas, passam por uma rotina sistemática de revisão, antes de serem adicionadas à base de dados. Periodicamente, novos documentos são adicionados e novas funcionalidades são integradas. Através da opção Fale Conosco, os usuários fazem contato com a administração do site, reportando eventuais falhas, erros, dúvidas e

5 Disponível em: <www.lampeh.ufv.br/acervosmg>. Acesso em: 09 maio 2011.

sugestões. Sempre que possível e viável, são feitas alterações, de modo a incorporar sugestões e críticas dos usuários.

O sítio do projeto está *online* desde fevereiro de 2008, tendo passado por uma série de testes e aperfeiçoamentos de estrutura e *layout*, desde então. O sistema se apresenta estável e seguro, permitindo acesso rápido, com boa interatividade e alta qualidade de resolução de imagens. Contamos atualmente com aproximadamente 400 usuários cadastrados, de diversos estados da federação e do exterior, e o número de acessos é crescente.

As imagens digitalizadas estão sendo gradativamente incluídas na página, após cuidadosa revisão. Atualmente, estão disponíveis para acesso livre todos os inventários digitalizados entre 1713 e 1770, totalizando cerca de 360 autos e aproximadamente 29.000 imagens de alta resolução.

Figura 5 - Página inicial do sítio



6 Conclusões

No início do trabalho, foram definidos requisitos de várias naturezas (desempenho, segurança, atendimento aos padrões e acessibilidade) que o sistema deveria cumprir. A estrutura e as tecnologias utilizadas no desenvolvimento, foram selecionadas com o intuito de atender a esses requisitos. Como requisito complementar, procurou-se adotar padrões

abertos e tecnologias de software livre, de modo a não se prender a fornecedores específicos. Os metadados selecionados para o padrão de arquivamento seguiram os recomendados pela norma ISAD(G), no entanto, a exportação dos metadados no formato XML, associado à norma, ainda não foi implementada, mas está prevista para as próximas versões. Para atender ao requisito desempenho, o tamanho das imagens de documentos foram reduzidas para um tamanho que viabilizasse uma transmissão rápida, porém, ainda mantendo uma qualidade de imagem aceitável. Este requisito possui um ponto de equilíbrio difícil de ser atingido e, nesse caso, algum sacrifício deve ser feito em algum ponto. Particularmente, no caso do sistema desenvolvido, clientes com conexão lenta reportaram muita lentidão no acesso aos documentos. Futuras versões devem permitir que o usuário especifique a velocidade de sua conexão para que o sistema envie os documentos em um tamanho mais adequado, com possíveis perdas na qualidade da imagem. O requisito de segurança foi atendido por meio de uma solução simples de cadastramento e de aceitação do termo de compromisso quanto ao uso das imagens dos documentos. Uma marca d'água também foi adicionada às imagens, para identificar sua origem. Para atender ao requisito acessibilidade, foi implementada a possibilidade da navegação em uma árvore, seguindo a ordem cronológica, além da busca exata e aproximada sobre os metadados, onde o pesquisador pode selecionar os campos onde será realizada a busca. O usuário pode, ainda, acessar os instrumentos de pesquisa, que dá acesso a arrolamentos onomásticos e cronológicos, da totalidade do acervo físico de inventários da Casa Setecentista de Mariana. Nestes arrolamentos, reunidos no formato PDF, o usuário pode identificar documentos de interesse, através da localização do acervo físico ("caixa" e "auto") ou pelo nome do inventariado, do inventariante ou pela data.

Nem todos os inventários arrolados estão disponíveis atualmente no site. Para verificar a existência do documento digital, deve-se realizar uma busca exata, usando-se o nome do inventariado. Caso o documento ainda não esteja disponível, o consulente tem a opção de consultá-lo diretamente na Casa Setecentista de Mariana.

Como parte de um projeto paralelo, os documentos estão sendo transcritos, de modo que, no futuro, será possível fazer buscas no próprio texto do documento. O processo de transcrição é lento, uma vez que a transcrição exige prática e conhecimento sobre a forma de redação de cada período.

Conclui-se, então, que todos os requisitos foram atendidos, sendo que alguns em maior extensão do que outros. Alguns aperfeiçoamentos, detectados a partir do *feedback* dos usuários, necessitam ser realizados e devem ser incorporados ao sistema, nas próximas versões.

Considera-se, também, que as metas de conservação preventiva, organização, informatização do acervo, digitalização e acesso público, por meio digital, foram plenamente atingidas, tendo o projeto colaborado, de modo significativo, para a preservação e divulgação do importante acervo

histórico da Casa Setecentista de Mariana e para o Resgate da Memória Histórico-Cultural de Minas Gerais.

Referências

INTERNATIONAL COUNCIL ON ARCHIVES (ICA). *ISAD(G): General International Standard Archival*. 2. ed. Ottawa: CIA/CDS, 2000. Disponível em: <[http://www.icacds.org.uk/eng/ISAD\(G\).pdf](http://www.icacds.org.uk/eng/ISAD(G).pdf)>. Acesso em 12 set. 2007.

INTERNET IMAGING PROTOCOL (IIP). *Version 1.0.5*. [s.l.]: Eastman Kodak Company; Hewlett Packard Company; Live Picture, 1997. Disponível em: <<http://iipimage.sourceforge.net/IIPv105.pdf>>. Acesso em: 13 maio 2009.

GARRETT, J. J. *Ajax: a new approach to web applications*. 2005. Disponível em: <<http://www.adaptivepath.com/ideas/essays/archives/000385.php>>. Acesso em: 13 maio 2009.

MENDES, F. F. *et al.* Tecnologia a serviço da história. *Revista do Arquivo Público Mineiro*, v. 43, p. 162-167, 2007.

MOREIRA, A. *et al.* Digitalização de manuscritos históricos: a experiência da Casa Setecentista de Mariana. *Ciência da Informação*, Brasília, v. 36, n. 3, p. 89-98, 2007.

ZIVIANI, N. *Projeto de algoritmos com implementações em Pascal e C*. 2. ed. São Paulo: Pioneira Thomson Learning, 2004.