

Comentário VII

José Lavaquial Breitinger

DocPro - Criação de Bibliotecas Virtuais Ltda

O tema instigador apresentado por Ricardo Mendes tem sido nossa vivência diária nos últimos anos. Temos nos defrontado com várias linhas de pensamento, como não poderia deixar de ser em qualquer nova tecnologia. Não precisamos ir muito ao passado, com os carros não foi diferente. Provavelmente na época muitos guardaram suas diligências para “se faltar combustível”, pois muito tempo levou para se padronizar o carro como o conhecemos hoje.

Vamos tentar sintetizar nossas experiências e dar uma visão pragmática do que temos observado nas instituições em que colaboramos na disponibilização de suas obras históricas ou raras via imagens digitais tanto para uso interno quanto na Internet. Entre elas citam-se as obras raras da Biblioteca Nacional e da Mário de Andrade, do Museu Histórico Nacional, do Museu do Folclore, do Instituto do Patrimônio Histórico e Artístico Nacional (Iphan), da Academia Brasileira de Letras, da Casa de Rui, da Fundação Getúlio Vargas, etc.

Em algumas instituições as pessoas estão a princípio abertas, em outras tem conceitos predefinidos, normalmente por ter lido algum artigo, em especial, estrangeiro.

Afirmações como “a imagem tem que ser no mínimo de tantos dpi”, ou “tem que ser neste ou naquele formato” ou ainda “tem que ter versão reduzida para Internet e alta para arquivo” estão a princípio conceitualmente incorretas ou imprecisas. Já expressões “microfilme é para preservação e imagem digital para exibição” estão mais longe ainda da realidade.

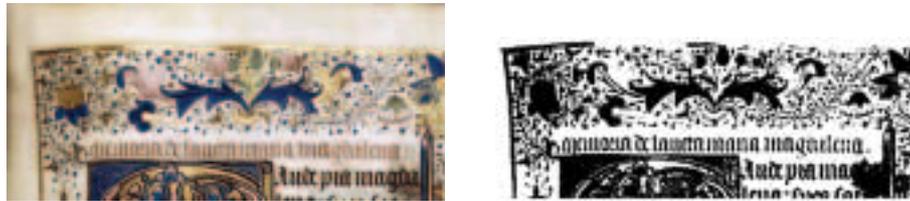
Por que são falhas essas argumentações?

A maioria delas baseia-se em fatos temporais, ou seja, como a tecnologia digital está em ebulição com alta taxa evolutiva, o limitante de hoje é facilmente realizável amanhã.

Sendo um pouco mais técnico, micros mais rápidos, portas de comunicação (USB, SCSI) mais velozes, HDs, DVDs e CDs com maior capacidade,

câmeras digitais com mais definição, scanners mais sofisticados, Internet banda larga, algoritmos de compressão, etc. mudam essa realidade a cada dia.

“A imagem tem que ser no mínimo de tantos dpis.” Aí vem a pergunta: estamos falando de cores, cinza ou PB (bitonal)? Uma imagem em cores precisará de menos dpis que sua equivalente cinza e menor ainda em relação a PB, para ter efeito melhor nas visualizações e impressões. A imagem PB perde detalhes que a colorida não perde. Nas imagens abaixo, qual sairá melhor impressa?



Livro de Horas, séc. XIV – Obras Raras BN.

“Imagens reduzidas para Internet” – as obras raras da BN, da Mário, etc. estão em alta definição (200 dpis ou mais a maioria), em cores, na íntegra, na Internet. A resposta é tecnologia, pois o usuário, dependendo da sua velocidade na Internet, pode ver com maior ou menor qualidade. Essa redução é feita pelo *software* de exibição em tempo real, para cada imagem vista.

“A imagem tem que ser neste ou naquele formato.” O formato/compressão da imagem é um item de suma importância. Vemos algumas pessoas usando Tiff para imagens coloridas ou cinza. O padrão Tiff – compressão Ccitt IV é excelente para imagens PB, usá-lo em coloridas não faz muito sentido, pois ou se usa um padrão de compressão como o LZW ou a compressão vai ser mínima. O LZW é um risco, pois em breve deve sumir, por não ser muito usado e ter problemas com licença até pouco tempo atrás. Para coloridas, no momento o padrão JPG é o mais indicado. Sabendo suas limitações e seu uso correto, a perda de qualidade é imperceptível e o resultado excelente. Padrões novos estão surgindo (JPEG2000 ou Djavu), mas ainda não é o momento para usá-los em acervos perenes, pois hoje poucos *softwares* as lêem e não há certeza de sua aceitação nos próximos anos. Corre-se o risco de perder o acervo em pouco tempo.

○ que é efetivamente importante

Disponibilizar acervo. Acervos demonstrativos (uma ou outra página/foto, só as representativas, etc.) estão perdendo ou perderão espaço. Os acervos de obra completa (como os Tesouros da BN e da Mário de Andrade) ganharão esse espaço, disponibilizando os fac-símiles dos originais na íntegra com boa definição. Apresentação só dos catálogos com os descritores das obras também tenderá a ser insuficiente.

Infelizmente a disponibilização de acervo na íntegra é o assunto em que mais se tem patinado e são poucos os *sites* na Internet com conteúdo na íntegra. Claro que para disponibilizar grandes volumes de conteúdo é preciso ter tecnologia de visualização e pesquisa, dessa forma metas, custos e prazos podem se cumpridos. Instituições querendo desenvolver sua própria técnica normalmente subestimam os problemas e terminam com um acervo digital de qualidade ruim e com ferramentas de acesso e pesquisa primitivas, além de excederem prazos e custos quando conseguem chegar ao final. E, em instituições públicas é muito difícil ficar em dia com os avanços tecnológicos, por cauda dos orçamentos irregulares (computadores, servidores, *scanners*, máquinas fotográficas, espaço de armazenamento, *links* de Internet).

Não ficar dependente de tecnologia. Um ponto importante é escolher o formato de imagem. Deve ser um formato *perene*, de *uso generalizado* e que dê boa qualidade e compressão, levando-se em conta os recursos computacionais de hoje tanto no armazenamento quanto na consulta. Pode-se e deve-se projetar os recursos um pouco à frente. Por exemplo, banda larga na Internet, hoje a maioria está na faixa de 256K. Mas em uma projeção para daqui a dois anos, pense em 512K ou 1Mega. Com isso em mente, no momento as opções são o Tiff e o Jpeg.

Já o PDF pode se tornar um problema no futuro, pois para acervos de imagens tem baixa qualidade e compressão, além de baixa *performance* nas consultas e pesquisas. É de se prever que haverá um substituto melhor (lembra-se do Carta Certa ou Word Perfect?), e então a dificuldade será "ler" os arquivos em PDF. Quer um exemplo? Imagine um livro com 300 páginas coloridas, 1Mb cada em JPG comprimido. Teremos na melhor hipótese um PDF com 200 Mb. Quem vai baixar isso pela Internet?

Mídia de armazenamento. Utilizar uma mídia de ampla aceitação. Os CDs e recentemente os DVDs são a melhor escolha, principalmente por seu uso doméstico (música e vídeo). Com isso se garante uma grande "inércia" para a obsolescência, assegurando que teremos equipamentos de leitura por muito tempo. E, por definição, qualquer leitor de DVD lê também CDs. Ou seja, os CDs estarão por aí por muito tempo. Uso de fitas DAT ou assemelhadas é desaconselhado para arquivos perenes, por não terem a tal "inércia" de mercado. É fantasia pensar que a duração dos CDs e DVDs é pequena, os de boa qualidade e bem tratados poderão durar mais de 50 anos.

Já os HDs podem ser usados por seu baixíssimo custo, porém exigirão mais atenção quanto à sensibilidade e à baixa perenidade. Ou seja, se for usado só HD, aconselha-se ter umas duas cópias *off-line* e a cada um ou dois anos copiar seu conteúdo para novas unidades. Caro? Negativo, um HD de 120Gb que custa R\$ 400,00 equivale a 200 CDs e a 30 DVDs. Qual instituição não poderá dispor de R\$ 800 por ano? É só programar e organizar a área digital, não a deixar abandonada. *Veja o próximo tópico.*

Ter equipe e procedimentos especializados na guarda digital. As instituições têm bons profissionais, conhecimento, estrutura e metodologia para a guarda e disseminação/duplicação de meio papel e microfilme.

Mas para o meio digital, quais as que têm esse mesmo aparato? É mais caro? Pelo contrário, um DVD pode ser duplicado por poucos reais e nele cabem aproximadamente 60.000 páginas PB. Quanto custa duplicar 60.000 páginas de microfilme em material, tempo, revelação e equipamento? E xerox de 60.000 páginas? Em um HD cabe 1,5 milhão de páginas PB ou 60.000 a 150.000 páginas coloridas. E o custo é só R\$ 400,00.

Qual a equipe certa? Provavelmente o pessoal que hoje toma conta do meio papel e microfimes. O pessoal da informática pode e deve ajudar, mas não devem ser os responsáveis pela guarda dos objetos digitais. Normalmente eles não são especializados em imagem e seria difícil o treino em técnicas arquivísticas. É mais fácil treinar especialistas da área de acervo em conservação, duplicação e impressão de imagens digitais.

Qual o perfil do profissional? Primeiro, conhecer técnicas arquivísticas, as mesmas aplicadas a microfimes e papel. Segundo, ter mente aberta a novidades e um quê de pesquisador e autodidata. Terceiro, gostar de computadores, ser aquele indivíduo que gosta de ensinar aos outros, de "futucar" o computador, que tem computador em casa e vive trazendo novidades. Difícil encontrar esse perfil entre arquivistas e bibliotecários? Nem tanto. Temos visto muitos nas instituições que atenderiam a essas características.

Não ter versão "reduzida" para Internet. Isso é fatal, é a famosa duplicação de dados. Sempre que alimentar uma base, é preciso alimentar a(s) outra(s). A cada imagem capturada, ter de produzir uma cópia com baixa definição e alimentá-la em outro sistema normalmente não funciona. O certo é ter tecnologia.

Sistema de visualização e pesquisa competente. Adianta digitalizar e não disseminá-las de forma efetiva? Fica sem sentido, pois a preservação dos acervos originais e divulgação não acontecerão. Para isso, é preciso:

1. Obter uma visualização de obras completas, com capacidade para milhares de livros ou álbuns com 300 ou mais imagens cada. Em cores, em alta definição. Com "trava" opcional que não permita cópia ou impressão da totalidade ou parte do acervo.
2. Ir diretamente a uma página sem ter que carregar as anteriores, o que tornaria inviável a visualização na Internet. Só por isso o PDF já deveria ser descartado.
3. Conseguir uma velocidade ajustável ao perfil de conexão à Internet do usuário, ou seja, ter-se menor ou maior qualidade na visualização. O ideal seria poder "navegar" em baixa qualidade e ao se encontrar a página/foto desejada, vê-la em melhor qualidade.
4. Disponibilizar opções de visualização: *thumbs* (ícones), *zooms*, visões panorâmicas, etc.
5. Ter capacidade para anexar objetos digitais, como descrições catalográfica, vídeos, sons, outras imagens, arquivos word, PDF,

etc. a qualquer imagem do acervo. Imagine o quanto seria enriquecedor ter ao mesmo tempo fotos ou documentos de Getúlio Vargas e um filme ou som original, um depoimento gravado sobre o assunto, uma reportagem recente, etc.

Final e principalmente, a *pesquisa*. Uma questão que vemos com frequência nas instituições é a desatualização e outros problemas na catalogação. Além disso, o usuário leigo está acostumado a pesquisar nos *goggles* da Internet. O ideal não seria disponibilizar pesquisas pelo catálogo e também por textos presentes nas imagens das obras digitalizadas? Sem maiores comentários, sugerimos as pesquisas abaixo, após carregar o *plugin* necessário.

- Pesquisar por *Napoleão Bonaparte* no texto presente nas imagens dos periódicos raros da BN e encontrar 16 páginas sobre o assunto, podendo escolher a qualidade de visualização do original. Sugestão: na tela que se abrirá, use a seta vermelha direita para ir de ocorrência em ocorrência. Ao chegar à página desejada, dê *zoom* e/ou clique no botão "colorido" se quiser vê-la colorida em alta definição.
- Pesquisar por *Lusíadas* (BN), encontrando a primeira página da primeira edição (1572) pelo texto de sua imagem. Pode-se ler a obra na íntegra. Dica: veja-a em cores, clique no "colorido" no topo a direita da tela que abrirá.
- Pesquisar na Mário de Andrade imagens de *Rugendas*, *Grahan* e *Borget* em cores, com *thumbs* (ícones). Detalhe – pesquisa baseada nas descrições catalográficas das obras e não no texto das imagens. Dica: use o *zoom* a 100%.

Se você quiser ver outras obras e instituições, acesse <http://www.docpro.com.br/bibliotecas.htm>. Se quiser ver curiosidades sobre obras raras, acesse <http://www.LinKultural.com.br>