

Carolina Vanuchi

Rui Pedro Almeida



Digitalização: problemas, recomendações e boas práticas

Trabalho da disciplina de Preservação e Conservação

Licenciatura em Ciência da Informação 2010/11

Docente: Prof^a Maria Manuela Pinto

Faculdade de Letras da Universidade do Porto

Janeiro de 2011

Resumo

O objectivo deste trabalho é compreender o que é digitalização, os problemas e a sua boa prática. A digitalização é um processo pelo qual uma imagem ou sinal analógico é transformado em código digital e pode assim ser armazenado ou e transferido para outros suportes.

Este trabalho expõe detalhadamente o conceito de digitalização, quais os cuidados que devem ser tidos em consideração ao digitalizar, o problemas que poderão surgir, assim como o planeamento que deve ser efectuado antes do processo. Uma parte importantíssima deste trabalho é a contextualização na área da Ciência da Informação e o que pode um profissional de CI trazer de novo. Para isso foi necessário comparar a visão do gestor de informação à de uma empresa, por exemplo.

Com a execução do trabalho pudemos alicerçar conhecimentos e adquirir novos em relação a este tipo de projectos que ao comum utilizador de um *scanner* parece simples. Após este estudo conseguimos criar um diagrama completo do processo operacional da digitalização que poderá vir a ser uma ferramenta muito útil, facilitando a execução de projectos deste tipo no meio onde estivermos a exercer a função como, esperamos nós, profissionais de CI. Ficamos também a conhecer as ferramentas utilizadas e todos os pormenores técnicos associados.

Sumário

1.	Introdução.....	4
2.	O processo da digitalização.....	5
2.1	Breve introdução à digitalização.....	5
2.1.1	Definição Geral.....	5
2.1.2	Tipos de digitalização.....	6
2.1.3	Imagem digital – definição e atributos	7
2.2	Importância da preservação digital	9
2.2.1	Motivos	9
2.2.2	Vantagens e limitações	10
2.3	Contextos de aplicação	12
2.4	Planeamento e execução.....	13
3.	Recomendações e boas práticas.....	16
3.1	Projecto MINERVA e NARA	16
3.2	Algumas dicas para a digitalização.....	17
4.	Infra-estruturas tecnológicas.....	19
4.1	Infra-estruturas técnicas e equipamentos.....	19
	• Máquinas Fotográficas Digitais	20
	• Scanner planetário	21
	• Scanner convencional ou tipo flatbed	21
	• Scanner de alimentação automática (ADF)	22
4.2	Suportes de armazenamento.....	22
5.	Considerações Finais.....	24
6.	Referências Bibliográficas	25
7.	Anexos.....	26

1. Introdução

Num mundo cada vez mais “digital”, a transferência de suportes têm um papel fundamental uma vez que é através destas aplicações tecnológicas que se conseguem as melhores medidas para neutralizar os potenciais factores de degradação dos documentos, o manuseamento, e simultaneamente se garante a melhor forma de acesso e disponibilização ao património.

Existem 3 tipos de transferência de suporte: a microfilmagem (usada para a preservação a longo prazo), a fotografia (preservação de suportes não físicos como memórias ou momentos) e a digitalização.

Com este estudo pudemos observar esse processo que exige, por parte dos responsáveis, acções prévias de organização, descrição, catalogação, indexação e a preparação física da documentação para se realizar a captura, armazenamento, recuperação e disponibilização com a qualidade necessária e para se garantir o respeito pela integridade do documento e a recuperação da informação no novo suporte.

O sucesso das pesquisas está dependente da qualidade de todo este trabalho e nesse sentido a disponibilização *online* ainda vem tornar mais evidente a necessidade da total articulação das acções referidas.

A implementação destas tecnologias acarreta às organizações investimentos para satisfazer necessidades em matéria de: recursos humanos – exigência de profissionais especializados, e recursos materiais – exigência de novos espaços de armazenamento, espaços de consulta, bem como de equipamentos, pelo que se exige a definição de políticas e estabelecimento de regulamentos que permitam identificar claramente os seus direitos e deveres, para assim poderem garantir a boa gestão da preservação e acesso.

2. O processo da digitalização

2.1 Breve introdução à digitalização

2.1.1 Definição Geral

A digitalização é um processo que consiste na transferência de informação no formato analógico para formato digital, permitindo o armazenamento e manipulação através de um computador. A digitalização contém duas actividades paralelas: a **amostragem** e a **quantificação**. A amostragem consiste em recolher amostras periódicas de um sinal analógico. A quantificação consiste em atribuir um valor numérico para cada amostra colhida.

A qualidade do sinal digital dependerá de dois factores:

- a frequência de amostragem (**taxa de amostragem**): quanto mais larga ela for (ou seja, uma recolha de amostras em menores intervalos de tempo) mais o sinal digital será fiel ao original;
- o número de bits em que codificamos os valores (chamado **resolução**): trata-se, na realidade, do número dos diversos valores que uma amostra pode comportar. Quanto maior, melhor a qualidade.

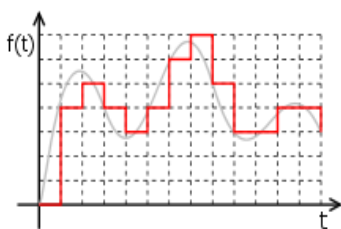


Imagem 1 - Sinal amostrado e quantizado

Assim, graças à digitalização pode-se garantir a qualidade de um sinal, ou então, reduzi-la para diminuir custos de armazenamento, digitalização ou o tempo de processamento. No entanto deve-se ter em conta o tipo de necessidade e as limitações materiais.

2.1.2 Tipos de digitalização

Ao contrário do que a maioria das pessoas pensam, quando se fala em digitalização de documentos não é somente de documentos em papel. Diversos suportes físicos e até mesmo abstractos como o caso do som podem ser digitalizados. Aqui apresentamos alguns exemplos de dados que podem ser digitalizados e de que maneira essa tarefa pode ser realizada:

- **Digitalização de documentos escritos em papel (informação textual):**

Passar documentos escritos em papel por um *scanner*, convertendo-os para o meio digital, é a aplicação mais directa e frequente da digitalização. Em geral os documentos em papel compõem a grande maioria dos repositórios das empresas (devido à facilidade de produção desses mesmos documentos), sendo eles também os principais responsáveis pela implantação de consultorias de gestão electrónica de documentos. Os documentos de papel possuem a vantagem da sua rápida produção mas, por outro lado, a sua recuperação e armazenamento são extremamente ineficazes e dispendiosos. Digitalizando os documentos do papel eliminamos estas duas desvantagens.

Ferramenta utilizadas: *Scanner* com OCR (Optical Character Recognition) ou ICR (Intelligent Character Recognition)

- **Digitalização de imagens (informação gráfica):**

Outro tipo muito comum e muito digitalizado são as imagens. Este tipo de suporte, se for armazenado de uma forma incorrecta, tende a deteriorar-se, enquanto as imagens digitais se preservam de modo indeterminado se forem tratadas devidamente, copiadas e armazenadas de maneira apropriada. Em imagens digitalizadas, também temos vantagem na organização e recuperação que só é possível com a digitalização destes documentos.

Ferramenta utilizadas: *Scanner* ou máquina fotográfica digital.

- **Digitalização de som e vídeo (informação audiovisual):**

O som individualmente ou aliado ao vídeo é um poderoso instrumento de trabalho e muitas vezes a informação é produzida e armazenada neste meio.

No entanto, muitas vezes é armazenado em suportes não adequados, de baixa qualidade e baixa durabilidade. Digitalizar o som e o vídeo significa um rigoroso tratamento ao nível do computador e devido armazenamento no formato digital mais adequado. Procedendo dessa maneira, além do ganho de qualidade ainda temos a vantagem da durabilidade dos dados que estarão melhor armazenados.

Chamadas telefónicas, músicas e filmes fazem parte de um grupo de informação em que a sua preservação é fundamental, por exemplo, para fins jurídicos.

Ferramenta utilizadas: *Scanner* ou máquina fotográfica digital.

2.1.3 Imagem digital – definição e atributos

A imagem digital é, por definição, é a representação de uma imagem bidimensional, sob a forma de uma grelha de píxeis coloridos, usando código binário de modo a permitir o seu armazenamento, transferência, impressão ou reprodução e processamento por meios electrónicos.

Existem várias formas de guardar imagens. Podem ser usados formatos em que é guardada a informação relativa a cada unidade mínima da imagem e formatos onde é guardada informação sobre a imagem sob a forma de vectores. Os formatos com informação sobre todos os pontos, quando são ampliados perdem qualidade, enquanto o formato da imagem vectorial é redimensionável sem perda de qualidade.

Os formatos de imagem têm muitas vezes opções que permitem indicar graus de compressão da imagem e outras especificações. Outro aspecto a reter é a resolução da imagem. Esta é medida em DPI (*dots per inch*) ou PPP (pontos por polegada). Seguidamente estão alguns valores considerados valores padrão:

- 72 dpi – adequado à visualização;
- 200 dpi – adequado à digitalização de documentos em sistemas de gestão documental;
- 300 dpi – qualidade usada para permitir a impressão à mesma escala e com a qualidade do original;
- 600 dpi – permite a impressão do documento ao dobro do tamanho original sem perda de qualidade;

Existem ainda formatos que permitem guardar várias páginas e formatos que permitem ter só uma página. Os formatos de imagem também necessitam espaços de armazenamento diferentes de acordo com a forma como armazenam a imagem.

Os formatos mais utilizados na digitalização são:

- **TIFF (*Tagged Image Format File*):**

Este é o formato mais utilizado para imagens digitalizadas. Este é também o formato utilizado pelos FAX, e tem diversas possibilidades de resolução. Os FAX são enviado com compressão do grupo 4 o que permite transformar uma página A4 em cerca de 30 Kb a 50 Kb. Este formato tem uma resolução de cerca de 200 DPI. Sem compressão uma página é pelo menos 10 vezes maior e se forem guardadas cores no TIFF o seu tamanho volta a multiplicar.

Recentemente foi criado o formato BigTIFF que pode ultrapassar o limite dos 4Gb do formato TIFF. A empresa canadiana Aperio, com este novo formato, criou a maior imagem digital do mundo com 143Gb e 1 terapíxel de resolução. A imagem trata-se de uma célula cancerígena.

- **PDF (*Portable Document Format*):**

Este formato é o mais utilizado para guardar em formato digital documentos que foram originalmente criados no computador. Ao guardar um ficheiro neste formato assegura-se que ele pode ser visto em diversos sistemas operativos e nos mais diversos tipos de computadores (como nos recentes Ipad ou Smartphones com Android OS). Nalguns casos os tipos de letra podem ser modificadas em relação às fontes originais.

- **JPEG (*Joint Photographic Experts Group*):**

Este é o formato mais utilizado e conhecido actualmente. A maior parte das máquinas fotográficas digitais guarda, por definição, as fotos em *.jpg* (extensão). É também muito utilizado na *Internet* e em multimédia, por ter uma excelente compressão, algo fundamental neste meio, e por suportar até 16 milhões cores distintas.

No entanto existem também outros formatos de imagem digital como: GIF, BMP, PNG, etc. mas que não são relevantes para esta temática.

2.2 Importância da preservação digital

2.2.1 Motivos

Por que razão devem ser digitalizados os documentos? Aqui encontram-se alguns motivos que mostram a importância deste processo:

Recuperação e selecção de documentos:

Os documentos não têm valor se não puderem ser encontrados. E se estiverem desorganizados, levarão mais tempo a ser encontrados e, conseqüentemente, essa desorganização será dispendiosa financeiramente por gastar tempo.

- **Protecção contra perdas:**

Os documentos importantes podem ser protegidos contra danos, erros no arquivamento ou até perdas;

- **Redução de custos:**

Um funcionário de escritório poderá gastar várias horas por semana a tentar localizar documentos e o custo do arquivamento de um documento em papel é geralmente alto se considerado o espaço utilizado. O arquivamento digital de documentos pode reduzir significativamente os custos relacionados dentro de uma organização;

- **Partilha eficaz de informação:**

Fazer grandes volumes de cópias de documentos em papel para partilhar a informação tem custo elevado tanto ao nível financeiro como ao nível do impacto ambiental. Documentos digitalizados podem ser enviados por e-mail ou disponibilizados instantaneamente via rede interna ou via Internet;

- **Preservação do conhecimento:**

Documentos em papel podem ser danificados facilmente e ficarem desgastados com o tempo. Digitalizados eles ficam protegidos das alterações climáticas, do passar dos anos e do manuseamento;

- **Segurança:**

Através da implementação de palavras-passe, somente pessoal autorizado tem acesso à documentação, reduzindo o risco de pessoal não-autorizado ter acesso a informação confidencial dirigida a grupos restritos;

- **Redução de riscos:**

Se um incêndio ou uma inundação atingirem uma organização, os documentos em papel serão facilmente destruídos, resultando numa potencial perda de informação essencial. Uma vez que os documentos tenham sido armazenados digitalmente, podem ser facilmente transferidos para o formato desejado - como um CD, DVD, Disco Duro, Memória Flash - e armazenados seguramente noutra local. Documentos esses que podem também ser acedidos remotamente.

2.2.2 Vantagens e limitações

- **Economia de tempo e espaço:**

A documentação digitalizada pode ser acedida mais rapidamente pois não implica a deslocação física do material. A poupança de espaço físico que advém da digitalização permite economizar e libertar espaço nas infra-estruturas. O material digital é bastante mais barato no que toca à sua reprodução e disseminação através das modernas redes de computadores.

- **Preservação dos originais:**

A digitalização permite manter intacta a documentação original, garantindo a sua segurança e conservação.

- **Segurança e confidencialidade:**

Toda a documentação digitalizada é armazenada de acordo com rigorosos sistemas controlo de acessos e segurança. O manuseamento e o processo produtivo são executados de forma a não causar dano nas espécies documentais analógicas e digitais.

Embora tenha todas as vantagens enunciadas anteriormente, a digitalização de documentos tem algumas limitações. No entanto, todas elas podem ser devidamente evitadas, se a gestão dos documentos digitais for realizada por uma equipa especializada.

- **Limitação de tempo:**

Embora durem até 100 vezes mais, os suportes de armazenamento digital também têm limitação de tempo de duração. Mas esse limite vai depender da maneira como são conservados suportes como o CD, o DVD, etc. No entanto existe uma solução: como a cópia destes formatos não tem perdas de qualidade, pode ser feita esporadicamente e de forma rápida para contornar esse limite de tempo inevitável.

- **Limitação de espaço:**

Esta é também uma limitação óbvia. No entanto, além do baixo custo de investimento, o suporte digital tem outra vantagem: os documentos digitalizados podem ser compactados para poupar ainda mais espaço.

- **Limitação de reprodução:**

Muitas vezes um documento digital é armazenado num formato e alguns anos depois já não existe um programa específico para abri-lo. Este é um problema mais sério que os dois apresentados acima, mas também é passível de solução. Hoje em dia para evitar que este caso ocorra, as indústrias tanto de software, como de

hardware têm criado padrões (Mpeg Layer 3, JPG, USB, etc...) que são usados por múltiplos programas e máquinas para poder aumentar a compatibilidade. Isto garante que a digitalização seja feita apenas uma vez.

- **Limitação do suporte legal:**

Outra grande limitação da digitalização é a falta de suporte legal para a mesma. No entanto a curto prazo é uma limitação que irá deixar de existir visto estarem a ser reunidos esforços e iniciativas neste sentido.

- **Fragilidade e ciclo de vida dos suportes de armazenamento:**

Seja qual for o suporte digital utilizado tem sempre associado uma elevada fragilidade e um ciclo de vida curto em relação ao papel.

- **Evolução tecnológica:**

A conseqüente evolução do hardware e software utilizado na digitalização e na pós-digitalização, pressupõem um olhar atento às actualizações e um largo fundo de investimento.

2.3 Contextos de aplicação

Existem dois contextos onde a digitalização é uma acção essencial: Na gestão documental e na preservação digital. Muitas empresas têm dificuldade na administração, organização e gestão da informação gerada e mais de 50% da informação produzida nas organizações, quer seja estratégica ou de decisão de negócios, é armazenada em computadores pessoais ou portáteis que não têm o suporte ideal em relação ao armazenamento ou backup. Logo, uma boa gestão do ciclo de vida da informação torna-a muito mais produtiva ao longo desse mesmo ciclo.

Os canais utilizados para troca de informação mudaram. Contudo não devem constituir surpresa os métodos utilizados para capturar, gerir e guardar os dados. Nos negócios contínuos e em crescimento, pode existir a necessidade de armazenar e recuperar dados de diferentes fontes, como o e-mail, Office, chamadas de voz, documentos técnicos, faxes entre outros.

O multi-acesso (inclusive remoto), o fácil manuseamento e a maior rapidez na distribuição de conteúdos resultantes da implementação de projectos de digitalização contribuem significativamente para uma maior organização e eficácia quanto à produção, armazenamento e tratamento da informação.

Desta forma, para evoluir com a tecnologia, os consultores e integradores devem familiarizar-se com os factores que levam à captura de informação, bem como com os novos conceitos de integração que giram em torno deste assunto.

Já no contexto da preservação de informação enquanto profissionais de CI, o estudo da mesma apresenta-se como algo fundamental, em que a análise da evolução dos métodos, dos procedimentos e das maneiras de ver essa mesma preservação ao longo dos anos constituem uma grande parte dos projectos de investigação desenvolvidos. Não menos importante, todavia, são os esforços actualmente desenvolvidos em prol da «preservação digital», de uma preservação cada vez mais inseparável dos processos da gestão da informação, considerada um recurso de gestão organizacional e de memória individual e colectiva.

2.4 Planeamento e execução

Antes de se iniciar ao planeamento deve-se ter atenção aos seguintes pontos e equacionar os seguintes aspectos

- A necessidade de preservação da informação;
- A verba suficiente para a realização do projecto;
- O pessoal qualificado para o seu desenvolvimento e em número suficiente;
- Capacidade e infra-estrutura técnica para a realização do projecto;

É importante salientar que os métodos, procedimentos, equipamentos e materiais utilizados para a prestação do serviço de digitalização obedecem escrupulosamente a vários critérios. A salvaguarda da integridade intelectual e física dos documentos originais, a rapidez e segurança na execução, a fidelidade das reproduções digitais em relação aos respectivos originais, conformidade com as normas que regulam os equipamentos, materiais e procedimentos de reprodução de documentos assim como um controlo rigoroso de qualidade do processo de digitalização e do produto final fazem parte dos critérios a serem obedecidos.

Quanto ao planeamento, propriamente dito, é o primeiro passo em qualquer projecto de digitalização. O tempo despendido no planeamento do projecto será sempre recompensado na simplificação da gestão e execução do projecto. Normalmente, as seguintes questões terão de ser respondidas:

- Que trabalho tem de ser feito?
- Quem deverá fazê-lo?
- Onde deverá ter lugar?
- Quando terá lugar?
- Como será feito?

Um projecto de digitalização deve ter metas e objectivos claramente definidos, pois estes terão impacto directo na selecção e na protecção dos direitos de autor e publicação.

O projecto deverá ter pessoal adequado, com conhecimentos e capacidades apropriadas, bem como um plano de formação para fornecer as especializações adicionais que o projecto possa requerer. Assim, é importante também que o pessoal necessário para trabalhar no projecto esteja disponível. Muitas instituições não têm grande quantidade de pessoal com muito tempo livre para dedicar a projectos de digitalização para além do seu trabalho habitual. Os conhecimentos necessários para projectos de digitalização podem também ser diferentes dos necessários para a realização das tarefas diárias. A equipa responsável terá de identificar soluções de *software* e de *hardware* específicas para um projecto de digitalização.

Na construção da equipa é preciso ter em conta os seguintes aspectos:

- Assegurar que existe a possibilidade de criar uma equipa com um número adequado de membros para realizar o projecto;
- Designar pessoas para cada tarefa ou partes do plano de projecto;
- Identificar necessidades de formação, incluindo formação na área de tecnologia da informação e educação no manuseamento de documentos delicados;
- Se possível, realizar a formação usando o *software* e *hardware* que irão ser utilizados durante o projecto, antes de o mesmo começar (os fornecedores

oferecem por vezes soluções técnicas gratuitas para formação ou poder-se-á alugar equipamento pretendido por um curto período de tempo).

- Optar por um núcleo pequeno de pessoal especializado dedicado ao projecto, em vez de um grande grupo de pessoal criado para o efeito.

Um projecto não deve ser iniciado antes de ser levada a cabo uma pesquisa de outros projectos na mesma área. Essa pesquisa vai identificar questões que têm de ser repensadas, vai estimular novas ideias aspectos que podem não ter sido muito explorados, já não mencionando o facto óbvio de que essa pesquisa prévia acrescentará valor e credibilidade aos resultados do projecto.

A pesquisa vai também ajudar a indicar a quantidade de trabalho que poderá ser planeado para a execução do projecto, através de contactos com entidades que tenham realizado projectos semelhantes. Estas interações irão ajudar a avaliar se existe o pessoal necessário, a capacidade e as infra-estruturas tecnológicas suficientes para avançar com o projecto. Se não houver pessoal especializado é necessário estudar uma forma de formar o pessoal existente.

Algum tempo pode ser investido, com rentabilidade, na verificação dos direitos de autor sobre o material a ser digitalizado. A falha em assegurar a permissão para digitalizar e publicar na Web pode ser a falha de um projecto de digitalização, independentemente da especialidade técnica em causa.

Deve ser também considerada a existência de um orientador técnico ou uma equipa, a envolver logo no início do projecto, que assegure que alguma anomalia ou problemas com a componente técnica sejam resolvidos antes do arranque do mesmo.

3. Recomendações e boas práticas

3.1 Projecto MINERVA e NARA

O projecto Minerva, criado em 2002 sob a direcção do Ministério da Cultura Italiano. O projecto incluiu representantes de ministérios de vários países da União Europeia, com o objectivo de promover abordagens e metodologias comuns para a digitalização de material cultural Europeu. O projecto reconheceu o valor único da herança cultural Europeia e o papel estratégico que a mesma pode desempenhar no crescimento da indústria de conteúdos digitais na Europa. Também reconheceu o valor da coordenação de esforços de governos nacionais e de organizações culturais, com vista ao aumento do nível de desenvolvimento e sinergia entre e dentro de iniciativas de digitalização.

Neste projecto foram apresentadas as mais importantes lições aprendidas e informação recolhida pela equipa responsável pelo projecto de boas práticas Minerva, com especial atenção para um número significativo de regras básicas que deverão ser consideradas por organizações que estejam a planear, executar ou gerir projectos de digitalizações no meio cultural. As orientações estão divididas nas seguintes áreas, cada uma reflectindo um nível no ciclo de vida de um projecto de digitalização:

- Planeamento do Projecto de Digitalização;
- Escolha de Materiais Originais para Digitalização;
- Preparação para a Digitalização;
- Manuseamento de Originais;
- O Processo de Digitalização;
- Preservação das Matrizes Digitais;
- Metadados;
- Publicação;
- Protecção da Propriedade Intelectual e Copyright;

- Gestão de Projecto.

Já nos EUA, originalmente, cada ramo e agência do governo dos Estados Unidos da América era responsável pelos seus próprios documentos e, infelizmente, por falta de conhecimento, muitas vezes existia perda e destruição de informação. Assim, o Congresso estabeleceu o Arquivo Nacional em 1934 com o fim de centralizar o cuidado dos arquivos. O Arquivo Nacional foi incorporado na Administração Geral de Serviços em 1949, mas, em 1985, foi feita agência independente, conhecida como NARA (National Archives and Records Administration).

A maioria dos documentos ao cuidado da NARA é de domínio público, visto que os trabalhos do governo estão excluídos da protecção de direitos de autor. A NARA também armazena documentos confidenciais.

3.2 Algumas dicas para a digitalização

Nesta secção deixamos uma série de dicas e conselhos para projectos de digitalização ou mesmo digitalizações a nível doméstico:

Digitalizar em *scanners* de mesa apenas material que não se estrague por ser pressionado sobre uma superfície plana e dura. Em caso de dúvida, consultar os especialistas.

Assegurar que o vidro do tabuleiro de digitalização está sempre totalmente limpo. Isto dará por um lado melhor qualidade de imagem, por outro protegerá o material original de sujidade. Se possível, digitalizar apenas itens que caibam por inteiro no *scanner* de mesa ou no scanner equipado com suporte de livros.

Se for necessário digitalizar um item em múltiplas partes, assegurar que se digitalizam margens de sobreposição suficientes para permitir a remontagem da imagem no pós-processamento (usando soluções de software de criação de mosaicos).

Testar o *scanner*, e os seus resultados, em material não-sensitivo antes de começar a digitalizar o material original. Os utilizadores devem também ser treinados com o mesmo material não-sensitivo.

Estabelecer uma convenção para a nomeação de ficheiros produzidos pelo *scanner*, usando por exemplo o sistema de catalogação existente ou dando-lhes

nomes significantes. O nome do ficheiro deverá permitir o estabelecimento da relação ficheiro-fonte original.

De maneira a maximizar a portabilidade de ficheiros entre plataformas de computadores, deve utilizar-se nomes de ficheiros com o máximo de oito caracteres, seguidos de uma extensão com o máximo de três caracteres.

Antes de estabelecer um fluxo de trabalho ou processos de articulação de trabalho, executar todo o processo de digitalização e processamento de imagem, de maneira a assegurar que o resultado final será o previsto.

Digitalizar com a máxima resolução e profundidade de bits que seja possível, de acordo com os motivos do projecto, as limitações do *scanner*, as condições de armazenamento de dados e os atributos do material original.

Digitalizar com a máxima profundidade de cor adequada, de acordo com as mesmas limitações.

Realizar *backups* (cópias de segurança) do disco rígido onde os dados estão armazenados, diariamente.

O controlo de qualidade da imagem digital e da meta informação é importante – o melhor momento para abordar questões de qualidade é no momento da digitalização.

Estabelecer parâmetros mínimos de resolução e cor (principalmente a resolução espacial e a profundidade de bits) para grupos de itens a ser digitalizados.

Verificar o resultado final do material digitalizado, no ecrã, no papel ou num outro formato a usar, por exemplo num dispositivo móvel.

Assegurar que os ecrãs (monitores) a ser utilizados estão bem calibrados. Evitar ter outro material em cima ou à volta do ecrã, pois pode afectar a percepção do item.

Devem ser criadas imagens para matriz com régua de escala, visíveis. Imagens coloridas ou em escala de cinzas, devem incluir um modelo *standard* de referência a cores ou com escala de cinzas.

4. Infra-estruturas tecnológicas

4.1 Infra-estruturas técnicas e equipamentos

Um projecto de digitalização tem de assentar sobre toda uma infra-estrutura técnica que não ponha em risco nenhuma fase do processo. No que concerne aos equipamentos, devem ser adequados ao tipo de informação que estiver a ser digitalizada. Esta é, de facto, uma operação relativamente simples. No entanto, de maneira a aumentar o nível de eficácia e eficiência e minimizar erros, será vantajoso ter um *workflow* operacional. A digitalização de itens excessivamente grandes, ou digitalização de alta qualidade, significa um grande investimento em tempo e esforço. Estes poderão ser reduzidos se se utilizar *hardware* apropriado ao documento. Não deverá, de todo, ser dispensada a presença de pessoas com formação.

Primeiro é necessário escolher bem o local de trabalho. É muito importante ter em conta aspectos como a refrigeração, tomadas eléctricas e de rede e, obviamente, o equipamento adequado. Estas decisões iniciais devem ser tomadas de acordo com as características do material que vamos digitalizar. O estado de conservação, as características físicas, o tipo de documento e a finalidade do ficheiro digital são alguns exemplos.

Quanto aos equipamentos que podem ser utilizados, temos:

- **Leitores Ópticos**



Ilustração 1 - Leitor óptico

Os leitores ópticos permitem ler códigos de barras ou outras indicações específicas e são muito usados na área comercial. Existem programas que permitem gerar códigos de barras específicos e imprimir esses códigos em etiquetas. O uso de

código de barras reduz o número de erros de introdução de dados de 1 em 300 para 1 em 3 milhões. Uma das empresas especialistas em Portugal em sistemas de leitura de códigos de barras é a Burótica.

- **Máquinas Fotográficas Digitais**



Ilustração 2 - Máquina fotográfica digital

A utilização de máquinas fotográficas digitais está a tornar-se muito comum em projectos de digitalização. Isto afirma a sua flexibilidade em termos de capacidade de digitalização de objectos que não sejam planos, assim como livros encadernados, manuscritos dobrados, e objectos tridimensionais.

Aqui estão apresentadas algumas indicações para o uso de máquinas fotográficas digitais na digitalização:

- Considerar a hipótese do aluguer de uma máquina de alta qualidade;
- Montar a máquina digital num suporte móvel apoiado numa coluna, e colocar os itens a ser digitalizados num suporte de livros estável com luzes adequadas;
- Solicitar ajuda/formação de um fotógrafo digital especialista – as diferenças de qualidade entre imagens tiradas por um amador e as tiradas por um especialista podem ser bastante significativas;
- Assegurar a utilização de um fundo que permite uma visualização clara do item;
- Evitar alterar as condições de luminosidade entre disparos, e entre imagens de diferentes partes de um mesmo item – o que poderá produzir impressões falsas de variações cromáticas.
- Usar lentes e filtros apropriados para combater problemas de registo de cor e distorção de imagem.

- **Scanner planetário**



Ilustração 3 - Scanner planetário

Este tipo de equipamento utiliza uma unidade de captura semelhante a uma máquina fotográfica digital, uma mesa de reprodução (que define a área de digitalização) e uma fonte de luz. São usados para a digitalização de documentos planos em folha simples, de documentos encadernados que necessitem de compensação de lombada (de forma a garantir a sua integridade física) e de documentos fisicamente frágeis, já que não ocorre nenhuma pressão ou tracção mecânica sobre os mesmos.

- **Scanner convencional ou tipo *flatbed***



Ilustração 4 - Scanner convencional

Este periférico pode ser encontrado facilmente em nossas casas. Acoplado a uma impressora multifunções ou como um aparelho individual, permite digitalizações domésticas ou não profissionais com uma qualidade adequada à necessidade

exigida. Alguns modelos (como é o caso do Scanner HP G2710, que tivemos o prazer de experimentar) possuem suporte para negativos e slides.

- **Scanner de alimentação automática (ADF)**



Ilustração 5 - Scanner ADF

A utilização destes equipamentos deve ser avaliada pormenorizadamente, devido ao risco potencial de danos físicos a comprometimento da longevidade dos mesmos, em virtude do seu modo de operar. Os equipamentos a serem utilizados devem ser adequados ao tipo de documento original, considerando principalmente os valores intrínsecos atribuídos ao documento: o seu estado de conservação e as suas características físico-químicas (dimensão, peso do papel, tipo de papel, tipo de tinta).

Todos os documentos a serem digitalizados principalmente neste tipo de equipamento deverão obrigatoriamente passar por um processo de análise rigorosa de sua estrutura física como por exemplo a remoção de sujidade e objectos como clips, grampos, fitas adesivas e afins.

Em anexo encontra-se um vídeo que mostra o funcionamento de um scanner ADF.

4.2 Suportes de armazenamento

Para garantir o sucesso a longo prazo de um projecto de digitalização, o produto (documento digital) tem de ser armazenado em suportes adequados e fiáveis. Existem dois tipos de suporte de armazenamento:

- Óptico (CD's, DVD's, Blu-Ray's, discos ópticos em ouro, discos ópticos em prata);
- Magnético (HDD's, Tape).

Estes tipos de armazenamento são classificados quanto à sua qualidade e durabilidade:

A seguinte tabela compara os suportes de armazenamento referidos anteriormente:

Suporte	Qualidade/Durabilidade
CD < DVD < Blu-ray	30 anos Diferença no número de camadas
Disco óptico em ouro	300 anos Resistente à degradação ambiental
Disco óptico em prata	100 anos maior resistência ao calor e aos raios UV
HDD	Fácil recuperação de informação e disponibilização automática
Tape	5 a 10 anos mais usado para Backups

5. Considerações Finais

Vivemos numa comunidade com milhares de anos em que a produção de informação é constante. Para evitar uma amnésia cultural, é necessário decidir como e onde armazenar toda essa informação.

A transferência de suportes, designadamente a digitalização, contribui para dois factores importantíssimos:

- A preservação da informação
- O aumento do acesso à informação

Todas as iniciativas que visem estes dois objectivos devem ser acompanhadas e abraçadas com o maior empenho possível pois delas depende toda a herança cultural que deixaremos às gerações seguintes.

A digitalização é um processo fácil, mas complexo pois carece das capacidades, cuidados, e preocupações que um profissional de CI deve estar dotado.

Com este trabalho pudemos penetrar no mundo da digitalização de uma maneira atenta, crítica, e absorvente e estamos em condições de concluir que acabou por ser uma investigação produtiva, aliciante e que nos irá trazer vantagens não só na nossa prestação em Preservação e Conservação, mas certamente para a nossa vida profissional.

6. Referências Bibliográficas

- Frequently Asked Questions About Optical Media. Disponível em: <<http://www.archives.gov/records-mgmt/faqs/optical.html>>
Consultado em: 23/11/2010
- INDIA. National Mission for Manuscripts. Guidelines for Digitization of Manuscripts. [Nova Deli], 2005. Disponível em: <<http://namami.org/digitization.pdf>>.
Consultado em: 24/11/2010.
- LIBRARY OF CONGRESS (Estados Unidos). Digital Preservation. Disponível em: <<http://www.digitalpreservation.gov/formats/fdd/descriptions.shtml>>.
Consultado em: 12/11/2010.
- LIBRARY OF CONGRESS (Estados Unidos). Formats, Evaluation Factors, and Relationships. Disponível em: <<http://www.digitalpreservation.gov/formats/intro/intro.shtml>>.
Consultado em: 12/11/2010.
- NATIONAL ARCHIVES AND RECORDS ADMINISTRATION (Estados Unidos). Electronic Access Project Scanning and File Format Matrix. Maryland, 1998. Disponível em: <<http://www.archives.gov/preservation/technical/guidelines-matrix.pdf>>. Consultado em 13/11/2010
- PROENÇA, Ana Luísa Morão Raposo Martins. Digital Preservation. Disponível em : <http://www.di.ubi.pt/~api/digital_preservation.pdf>.
Consultado em: 16/11/2010.
- PROJECTO MINERVA (USA) Disponível em: <<http://www.minervaeurope.org>>.
Consultado em 28/11/2010.
- SILVEIRA, Alex da. A informação na web: memória, digitalização, preservação e a importância do meio. Disponível em: < <http://bibliotecno.com.br/?p=632> >.
Consultado em 22/09/2010
- SONA, Creative Studio. Disponível em: <<http://www.sonna.com.br/digitalizacao.htm>> Consultado em: 01/12/2010
- TASI: Technical Advisory Service for Images. Disponível em: <<http://www.tasi.ac.uk/>> Consultado em: 14/11/2010

7. Anexos

- Anexo A – Processo da digitalização



- Anexo B –

[Vídeo de um scanner ADF em funcionamento](#)