

Instruções para regulamentação das técnicas de digitalização de documentos em suporte de papel

1. Objecto e finalidade

1.1. As “Instruções para regulamentação das técnicas de digitalização de documentos em suporte de papel”, doravante designadas por Instruções, estabelecem as exigências básicas e as normas orientadoras a observar pelos serviços públicos da Região Administrativa Especial de Macau, doravante designada por RAEM, na execução dos trabalhos de digitalização de documentos em suporte de papel, abrangendo as principais etapas do trabalho, as exigências de gestão e os critérios técnicos.

1.2. As Instruções visam criar regulamentação dos trabalhos de digitalização de documentos em suporte de papel, na qual inclui a do fluxograma dos trabalhos de digitalização, dos critérios técnicos e do formato das imagens, em benefício do controlo da qualidade dos trabalhos de digitalização, do intercâmbio de informações e da criação e acumulação dos recursos de arquivo do Governo da RAEM.

2. Âmbito de aplicação

2.1. As Instruções aplicam-se à digitalização de documentos em suporte de papel efectuada pelos serviços públicos da RAEM, nas seguintes situações:

2.1.1. Digitalização dos documentos em suporte de papel para preservação permanente, de acordo com as necessidades dos serviços.

2.1.2. Introdução de documentos em suporte de papel digitalizados nos sistemas de funcionamento dos serviços, para apoio e desenvolvimento de actividades quotidianas.

2.2. As Instruções não se aplicam no tratamento de documentos electrónicos originários.

3. Normas de referência

As Instruções têm por referência as seguintes normas:

3.1. GB/T 20530-2006 - Directrizes de digitalização de documentos e informações (norma nacional da República Popular da China).

3.2. DA/T 31-2005 - Regulamentação das técnicas de digitalização de documentos em suporte de papel (norma da indústria de arquivo da República Popular

da China).

3.3. ISO/TR 13028:2010 - *Information and documentation – Implementation guidelines for digitization of records.*

3.4. ISO/TR 15801:2009 - *Document management – Information stored electronically – Recommendations for trustworthiness and reliability.*

4. Termos técnicos

4.1. Digitalização (*digitization*)

Processo de tratamento de documentos mediante técnicas informáticas, através das quais os sinais analógicos são transformados em códigos digitais.

4.2. Digitalização de documentos em suporte de papel (*digitization of paper-based records*)

Processo de tratamento de documentos em suporte de papel através de equipamentos como o *scanner* e a câmara digital, transformando-os em imagens digitais identificáveis pelo computador e passíveis de serem arquivadas em suportes informáticos, designadamente em disco rígido, fita magnética, disquete e disco óptico.

4.3. Imagem digital (*digital image*)

Imagem resultante da digitalização de documentos em suporte de papel.

4.4. Imagem binária (*binary image*)

Imagem digital em preto e branco, com dois tons de cinza.

4.5. Imagem fixa com tom contínuo (*continuous-tone still image*)

Imagem fixa digitalizada, constituída com diferentes tons de graduação com mais de dois tons de cinza ou a cores.

4.6. Resolução (*resolution*)

Número de pontos e *pixels* que compõem a imagem digital numa determinada unidade de comprimento, de um modo geral pontos por polegada (*dpi*).

4.7. Medida de distorção (*distortion measure*)

Nível de desvio de cores e geometria da imagem digital relativamente ao documento original em suporte de papel sob o mesmo ambiente de ensaio, após a digitalização.

4.8. Inteligibilidade (*intelligibility*)

Capacidade de fornecer mensagens acessíveis a utilizadores e a equipamentos através de imagens digitais.

4.9. Compressão de imagem (*image compression*)

Remoção de redundâncias ou tratamento de semelhanças existentes na imagem

digital, de modo a representá-la de forma compacta.

5. Princípios a observar durante a digitalização de documentos em suporte de papel

5.1. Princípio geral

Durante os trabalhos de digitalização de documentos em suporte de papel devem escolher-se, de acordo com as situações previstas no ponto 2.1., os critérios técnicos mais adequados e satisfazer as exigências básicas da digitalização de documentos, de forma a permitir uma utilização correcta, cómoda, rápida e segura dos recursos de arquivo.

5.2. Princípio de avaliação

Antes de planear e iniciar os trabalhos de digitalização de documentos em suporte de papel deve proceder-se à avaliação da viabilidade e dos riscos da digitalização, ponderando o estado dos mesmos, os recursos e os equipamentos necessários.

5.3. Princípio de identificação

Antes de planear e iniciar os trabalhos de digitalização de documentos em suporte de papel deve proceder-se à classificação e identificação do valor dos mesmos, a fim de determinar, de entre as situações previstas no ponto 2.1., qual a mais adequada.

5.4. Princípio de legalidade

Os trabalhos de digitalização devem estar em conformidade com as disposições legais em vigor na RAEM.

6. Exigências básicas da digitalização dos documentos em suporte de papel

6.1. Planeamento do trabalho

Consoante as situações referidas no ponto 2.1., deve proceder-se ao planeamento tendo em conta o fluxograma dos trabalhos de digitalização, os suportes de arquivo das imagens digitalizadas e as especificações técnicas, bem como o planeamento do tempo, dos recursos humanos e dos custos necessários para efectuar a digitalização, a fim de formar um plano completo.

6.2. Etapas básicas

As etapas básicas da digitalização de documentos em suporte de papel são principalmente: a organização dos documentos, a criação da base de dados de índice, a realização do *scan*, o tratamento das imagens, o arquivo das imagens, a verificação

da qualidade dos dados, a interligação dos dados, o exame e a recepção dos dados, a execução de *backups* dos dados e a gestão dos resultados.

6.3. Gestão do processo

É necessário proceder ao registo de todos os trabalhos efectuados durante as diferentes etapas do processo de digitalização, de forma a obter um registo completo do mesmo.

6.4. Medidas de segurança

6.4.1. Para garantir a segurança do conteúdo dos documentos originais e dos digitalizados, é necessário estabelecer e aperfeiçoar mecanismos de gestão de segurança e de preservação da confidencialidade das diferentes etapas do processo de digitalização.

6.4.2. Os equipamentos afectos aos trabalhos de digitalização de documentos em suporte de papel devem estar, preferencialmente, isolados do exterior.

7. Organização dos documentos

Para assegurar a qualidade dos trabalhos de digitalização, antes da digitalização dos documentos em suporte de papel, estes devem ser organizados de acordo com as seguintes etapas:

7.1. Remoção da encadernação

Quando a encadernação afecte os trabalhos de *scan* ou prejudique a respectiva qualidade deve ser removida, devendo, durante a remoção, proteger-se o documento.

7.2. Restauro de folhas de papel

7.2.1. Os documentos gravemente danificados e que não possam ser digitalizados devem ser objecto de restauro técnico prévio.

7.2.2. Os documentos com folhas amarrotadas ou amassadas, que prejudicam a qualidade do *scan*, devem ser devidamente tratados, aplanando ou engomando as respectivas folhas antes de efectuar o *scan*.

7.3. Reposição da encadernação

7.3.1. Depois de efectuado o *scan*, o documento ao qual foi removida a encadernação deve ser novamente encadernado, de acordo com as exigências de arquivo de documentos.

7.3.2. Durante o processo de encadernação devem manter-se as folhas do documento na mesma ordem sequencial, assegurando o respectivo tratamento de forma segura, precisa e sem omissões.

8. Criação de base de dados de índice

8.1. Objectivo

A criação da base de dados de índice visa reforçar a gestão das imagens digitais a fim de facilitar a sua utilização futura.

8.2. Selecção do formato de dados

Aquando da criação da base de dados de índice, deve seleccionar-se o formato de dados mais comum e que permita a troca de dados, directa ou indirectamente, através de ficheiros XML.

8.3. Introdução de dados

8.3.1. Os documentos digitalizados devem ser registados na base de dados de índice.

8.3.2. Em relação aos documentos em suporte de papel ainda não digitalizados, devem ser introduzidos na base de dados de índice os respectivos dados de descrição, de acordo com a regulamentação definida sobre introdução de dados, com vista a facilitar a interligação e a utilização futuras.

8.4. Verificação da qualidade

8.4.1. Deve proceder-se à verificação da qualidade da base de dados de índice criada, da integridade dos itens de descrição, da conformidade do conteúdo com as exigências estabelecidas e da correcção do mesmo.

8.4.2. A verificação referida no ponto anterior pode ser efectuada através da revisão manual ou automática mediante programas.

8.4.3. Caso se verifique a existência de dados incorrectos, deve proceder-se às alterações necessárias ou a nova introdução.

9. Realização do *scan* de documentos

9.1. Forma de *scan*

9.1.1. Deve escolher-se o equipamento de *scan* com especificações adequadas, de acordo com o tamanho das folhas do documento original em suporte de papel.

9.1.2. Quando se trate de documento de grande dimensão, pode ser realizado o *scan* por partes, procedendo-se depois à costura das imagens, ou pode ser utilizada a captura fotográfica digital de imagens de grande dimensão.

9.1.3. Quando se trate de documentos em más condições, demasiado finos, macios ou grossos, a forma de *scan* a escolher deve evitar danos ao documento original.

9.1.4. Só pode adoptar-se a forma de *scan* de alta velocidade, a fim de elevar a eficiência dos trabalhos, para documentos em boas condições.

9.2. Modelos de cores de *scan*

9.2.1. Existem geralmente modelos de cores de *scan* a preto e branco, em tons de cinza e a cores.

9.2.2. Quando se trate de documento em suporte de papel com páginas pretas e brancas, com letras legíveis e sem figuras, deve utilizar-se o *scan* a preto e branco.

9.2.3. Quando se trate de documento em suporte de papel com páginas pretas e brancas, com letras pouco nítidas ou com figuras e letras em diversas cores, deve utilizar-se o *scan* em tons de cinza.

9.2.4. Quando se trate de documento em suporte de papel com páginas que contenham carimbos, fotografias a preto e branco, fotografias ou figuras a cores, deve utilizar-se o *scan* a cores.

9.3. Resolução do *scan*

9.3.1. A escolha da resolução não pode afectar a nitidez, a integridade e a utilização da imagem.

9.3.2. Nos casos especiais em que as letras ou caracteres sejam demasiado pequenos, densos ou pouco nítidos, pode aumentar-se adequadamente a resolução.

9.4. Registo do *scan*

Deve fazer-se o registo do *scan* do documento e do número de páginas, verificar se o mesmo coincide com o número de páginas do documento original e, em caso negativo, anotar o concreto motivo e a forma de tratamento.

10. Tratamento de imagens

10.1. Ajustamento da orientação

Rodar a imagem digital que não se encontre na orientação correcta para a posição habitual de leitura.

10.2. Correção de inclinação

Corrigir a imagem digital inclinada.

10.3. Eliminação de impurezas

Eliminar as impurezas que apareçam na imagem digital e que prejudiquem a sua qualidade, como manchas, linhas, enquadramento negro e bordas, etc.

10.4. Ajustamento da qualidade da imagem

Ajustar os parâmetros da imagem digital de qualidade insatisfatória após o *scan*, designadamente brilho, contraste, saturação, gradação de tonalidade, balanço de cores, etc., devendo nesse processo cumprir o princípio de mostrar o aspecto original do documento, com o pressuposto de não afectar a inteligibilidade do mesmo.

10.5. Costura de imagens

10.5.1. Quando o *scan* de documentos de grande dimensão é efectuado por partes, estas devem ter dimensões similares e a resolução e exposição devem ser as mesmas.

10.5.2. Devem existir sobreposição parcial entre as imagens digitais adjacentes e a área da sobreposição não deve ser inferior a 1/20 da área das partes.

10.5.3. Deve proceder-se à costura das várias imagens digitais criadas, formando uma imagem completa que assegure a respectiva integridade.

10.6. Corte da borda

Cortar a imagem digital para remover o espaço excedente, reduzir eficazmente o tamanho da imagem e poupar espaço de arquivo.

10.7. Autorização do tratamento de imagens e registo

A execução dos tratamentos de imagens referidos nos pontos 10.3. a 10.6. depende da autorização prévia do dirigente do serviço, devendo ser registada imediatamente e anotadas as causas concretas, o tratamento efectuado e as pessoas que o realizaram.

11. Arquivo de imagens

11.1. Formato de arquivo

11.1.1. As imagens digitais devem ser arquivadas em formato TIFF, devendo seleccionar-se o modo de não compressão ou de compressão sem perdas.

11.1.2. As imagens digitais disponíveis para consulta na rede também podem ser arquivadas em formatos PDF, JPEG ou outros.

11.2. Atribuição de nomes às imagens digitais

Cada documento em suporte de papel deve ter um único código, devendo, depois do *scan*, ser atribuído um nome à imagem digital com base nesse mesmo código e a denominação ser feita de forma uniforme.

12. Verificação da qualidade dos dados

12.1. Verificação da qualidade da base de dados de índice

Deve proceder-se à verificação da integridade do conteúdo da base de dados de índice e da conformidade do mesmo com a respectiva regulamentação, e à correcção ou a nova introdução dos dados incorrectos.

12.2. Verificação da qualidade das imagens digitais

12.2.1. Deve proceder-se a novo tratamento ou *scan* das imagens digitais com orientação, inclinação, nitidez e distorção que não correspondam às exigências de qualidade estabelecidas.

12.2.2. Quando se verifique falta de alguma página, deve proceder-se de imediato ao *scanda* mesma e introduzir-se correctamente a imagem digital omissa.

12.2.3. Se a ordem sequencial das imagens digitais produzidas não coincidir com a do documento original, deve proceder-se de imediato aos ajustamentos necessários.

12.3. Registo de verificação da qualidade

Deve fazer-se o registo de verificação da qualidade dos dados, indicar os resultados da verificação da qualidade e as sugestões de tratamento.

13. Interligação de dados

13.1. A fim de concretizar a interligação entre a base de dados de índice e as imagens digitais, a relação de correspondência deve ser estabelecida através da uniformidade e unicidade do código de documento da base de dados e do nome das imagens digitais.

13.2. Quando a base de dados de índice e as imagens digitais satisfaçam as exigências de qualidade estabelecidas, deve concretizar-se a interligação automática entre os dados de índice e as imagens digitais correspondentes, através da elaboração de programas ou através do recurso a programas adequados.

13.3. Deve fazer-se o registo da interligação de dados, inserindo o número de páginas obtidas depois da interligação e verificar-se se o mesmo coincide ou não com o número anotado durante o *scan*, em caso negativo, deve registar-se o concreto

motivo e a forma de tratamento.

14. Exame e recepção de dados

Se o trabalho de digitalização for executado por adjudicatário ou se existirem exigências específicas para o exame e recepção de dados, deve proceder-se nos seguintes termos:

14.1. Exame por amostragem dos dados

14.1.1. Deve efectuar-se, por amostragem, o exame da qualidade dos dados digitalizados.

14.1.2. A taxa de amostragem, por tipo de documento, que não deve ser inferior a 5%, é calculada de acordo com a fórmula seguinte:

$$\text{Taxa de amostragem} = \frac{\text{número de documentos da amostra}}{\text{número total de documentos a receber, por tipo}} \times 100\%$$

14.2. Critérios de exame e recepção

14.2.1. Caso a amostra da base de dados de índice e das imagens digitais não corresponda às exigências de qualidade estabelecidas ou a interligação entre a base de dados de índice e as imagens digitais seja defeituosa, será considerado reprovado o exame.

14.2.2. Se a taxa de aprovação dos dados da amostragem, por tipo de documento, for igual ou superior a 95%, é considerado aprovado o exame, podendo proceder-se à recepção do trabalho de digitalização.

14.2.3. Os dados não aprovados no exame são devolvidos à fase de verificação da qualidade, para verificação e correcção.

14.2.4. A taxa de aprovação referida no ponto 14.2.2. é calculada de acordo com a fórmula seguinte:

$$\text{Taxa de aprovação} = \frac{\text{número de documentos aprovados}}{\text{número total de documentos sujeitos a exame}} \times 100\%$$

14.3. Registo do exame e recepção de dados

Deve ser efectuado um registo detalhado do exame e recepção de dados, nomeadamente da taxa de aprovação, dos resultados do exame e recepção de dados, no caso de exames reprovados, dos motivos e do tratamento consequente.

15. Realização de *backups* (cópias de segurança de dados)

15.1. Âmbito dos *backups*

Deve proceder-se, em tempo oportuno, ao *backup* da base de dados de índice e das imagens digitais.

15.2. Formas dos *backups*

Para assegurar a segurança dos dados, a selecção dos suportes para proceder ao *backup* deve ser diversificada, podendo proceder-se a vários *backups*, mediante a combinação de *backups* acessíveis *online* e *offline*, sendo permitido, quando se revele necessário, efectuar o *offsite backup*.

15.3. Verificação de dados de *backup*

Devem ser efectuadas verificações periódicas dos dados de *backup*, as quais incluem principalmente o acesso aos dados de *backup*, a integridade das informações e a exactidão do número de imagens digitais.

15.4. Etiquetas de *backups*

Depois do *backup* de dados deve efectuar-se a etiquetagem nos meios de *backup* correspondentes, a fim de facilitar a procura, consulta e gestão.

15.5. Registo de *backups*

Deve fazer-se o registo dos *backups* de dados, indicando a quantidade de *backups*, a data de execução e o lugar de arquivo.

16. Gestão de resultados

16.1. Deve reforçar-se a gestão dos resultados da digitalização de documentos em suporte de papel, garantindo a segurança, a integridade e o acesso a longo prazo aos dados.

16.2. Para assegurar o acesso a longo prazo aos dados devem definir-se políticas de *backup* de dados, cujo conteúdo abranja os metadados correspondentes.

16.3. O prazo de conservação dos dados deve ser idêntico ao dos documentos originais a que correspondem.

16.4. O tratamento dos dados deve ser executado de acordo com os procedimentos de tratamento dos documentos em suporte de papel.

16.5. Quando os documentos digitalizados estiverem disponíveis para pesquisa e utilização em rede, pode, consoante a situação concreta, inserir-se identificação nas imagens digitais, e delimitar-se os *downloads* ou as impressões.

17. Critérios técnicos

17.1. Nas situações previstas no ponto 2.1.1. devem ser adoptados os critérios técnicos do Quadro I, sendo permitida a criação de réplicas com qualidade igual ou inferior a tais critérios.

Quadro I

Modelo de cores de <i>scan</i>	Resolução	Profundidade da cor	Formato	Compressão
Preto e branco	mínima de 300 <i>dpi</i>	1 <i>bit</i>	TIFF	compressão sem perdas
Tons de cinza	mínima de 300 <i>dpi</i>	8 <i>bit</i> escala de cinza	TIFF	compressão sem perdas
Cores	mínima de 300 <i>dpi</i>	mínima de 24 <i>bit</i> a cores	TIFF	compressão sem perdas

17.2. Nas situações previstas no ponto 2.1.2. devem ser adoptados os critérios técnicos do Quadro II.

Quadro II

Modelo de cores de <i>scan</i>	Resolução	Profundidade da cor	Formato
Preto e branco	mínima de 200 <i>dpi</i>	1 <i>bit</i>	-TIFF -PDF -JPEG
Tons de cinza	mínima de 200 <i>dpi</i>	8 <i>bit</i> escala de cinza	-TIFF -PDF -JPEG
Cores	mínima de 200 <i>dpi</i>	mínima de 8 <i>bit</i> a cores	-TIFF -PDF -JPEG