



Recomendações para a produção de
PLANOS DE PRESERVAÇÃO DIGITAL

Lisboa, 25 de novembro de 2011

Ficha técnica MIP

Título: Recomendações para a produção de Planos de Preservação Digital

Autor: Francisco Barbedo - coordenador

Autor: Luís Corujo

Autor: Mário Sant'Ana

Classificação: 260.01.01 DSIPE\001530 Vol01 Manuais técnicos área dos documentos eletrónicos

Descritores: Preservação Digital, Planos de Preservação Digital, Documentos de Arquivo Eletrónicos

Data/Hora: 2011-11-25

Formato de dados: Texto, PDF

Estatuto de utilização: Acesso público

Relação: versão - 2.1

© DGARQ, 2011

Alterações feitas a este documento

#	Alteração	Responsável	Versão
1	Alteração da nomenclatura de documento de arquivo Eletrónicos para informação de arquivo eletrónica (ou digital)	FB, LC, MS	V2.0
2	Criação de modelo de articulação entre diferentes entidades: informação de arquivo, sistema de informação, documento de arquivo Eletrónicos, objeto digital	FB	V2.0
3	Novos modelos de fichas de recolha de dados nos anexos	LC, MS	V2.0
4	Exemplos de aplicação adicionais incluindo modelos de preenchimento das fichas de recolha de dados nos anexos	LC	V2.0
5	Renumeração das fases de desenvolvimento de um PPD, contendo explicações adicionais	LC, MS	V2.0
6	Reposicionamento das bases de dados relativamente à estrutura hierárquica clássica de organização arquivística	FB	V2.0
7	Alteração do glossário para inserção de mais conceitos	LC	V2.0
8	Novos desenhos explicativos	FB	V2.0
9	Alteração de localização de definições de tipos de backup para o Glossário	LC	V2.0
10	Alteração da fonte para Trebuchet MS	MS	V2.1
11	Revisão do texto para a grafia determinada pelo Acordo Ortográfico da Língua Portuguesa de 1990	MS	V2.1
12	Alteração do logotipo	LS	V2.1

0	Introdução	5
0.1	A necessidade de um Plano de Preservação Digital (PPD)	6
0.2	Estado da situação.....	7
0.3	Preservação digital	8
0.4	O Plano de Preservação Digital	8
0.5	Objetivos e destinatários deste documento	9
0.6	Estrutura e conteúdo do documento	9
1	Elaboração de um Plano de Preservação Digital.....	12
1.1	Pré-requisitos.....	12
1.2	Metodologia para a elaboração do PPD	16
1.2.1	- Fase 1: Apresentação de metodologia para elaboração de um PPD	17
1.2.2	- Fase 2: Identificação dos sistemas de informação que vão ser objeto de PPD.....	17
1.2.3	- Fase 3: Avaliação arquivística dos sistemas de informação.....	19
1.2.4	- Fase 4: Caracterização tecnológica dos sistemas de informação que vão ser objeto de PPD e sua Avaliação	20
1.2.5	- Fase 5: Planeamento da estratégia de preservação	21
2	Implementação do Plano de Preservação Digital.....	26
2.1	Teste da solução proposta.....	26
2.2	Produção de cópias de segurança	26
2.3	Aplicação da Estratégia de Preservação.....	27
2.4	Medidas de monitorização e auditorias internas	27
2.5	Afetação de recursos	28
3	Ameaças	29
3.1	Obsolescência de formatos e software	29
3.2	Obsolescência do hardware e suportes	30
3.3	Elementos que afetam os suportes.....	31
4	Caracterização de formatos e aplicações	32
4.1	Formatos	32
4.1.1	Identificação de formatos alternativos normalizados	36
4.2	Aplicações.....	39
5	Sistemas de armazenamento	40
5.1	Atributos desejáveis	40
5.2	Arquiteturas Possíveis	41
5.2.1	Critérios para avaliação de soluções de armazenamento	43
5.3	Cópias de segurança.....	44
5.3.1	Gestão do repositório de dados.....	46
6	Metainformação	47
7	Estratégias de Preservação mais comuns.....	52
7.1	Preservação de tecnologia.....	52
7.2	Emulação.....	52
7.3	Monitorização de suportes e formatos	53
7.4	Encapsulamento.....	53
7.5	Transposição de Formatos e Suportes (Migração e transferência de suporte) ..	53
8	Conclusão	56
Anexos		57
Instruções de utilização do Anexo A - Lista de requisitos para a prática de preservação digital		58
A	Lista de requisitos para a prática de preservação digital	66
B	Folha de Recolha de Dados	70
Exemplo de preenchimento da Folha de Recolha de Dados		85
C	Exemplo de um ficheiro PREMIS contendo MI técnica NISO Z39.87	92
D	Recursos Aconselhados	95
E	Glossário.....	101

0 Introdução

Conceito - Ao longo deste documento serão utilizados os termos “documento de arquivo Eletrónicos” (DAE) “objeto digital” (OD) e “sistema de informação” (SI). Estes termos representam realidades relacionadas mas neste documento referidos com propósitos diferentes.

O DAE é a entidade lógica que possui conteúdo, contexto e estrutura de forma a ter um significado específico. Um sistema de informação é uma estrutura aplicacional especializada na contenção e gestão de dados e/ou informação. O OD é a componente física do DAE ou do SI, normalmente equivalente a ficheiros. Um DAE ou um SI é sempre composto por, pelo menos, um OD, podendo ser composto por vários OD. Por exemplo, um relatório em formato <.doc> pode ser composto por um ficheiro de texto, um ficheiro de imagem e um ficheiro tabular, estando todos integrados para a representação completa do documento. A mesma situação se aplica a documentos estruturados como bases de dados (simples ou de suporte a sistemas de informação complexos) ou ainda documentos multimédia.

Em termos práticos apenas se consegue preservar um documento de arquivo Eletrónicos se forem realizadas ações sobre as suas componentes, ou seja, os OD.

Noutra perspetiva e dada a pluralidade das representações de informação de arquivo atualmente existentes assume-se a seguinte categorização:

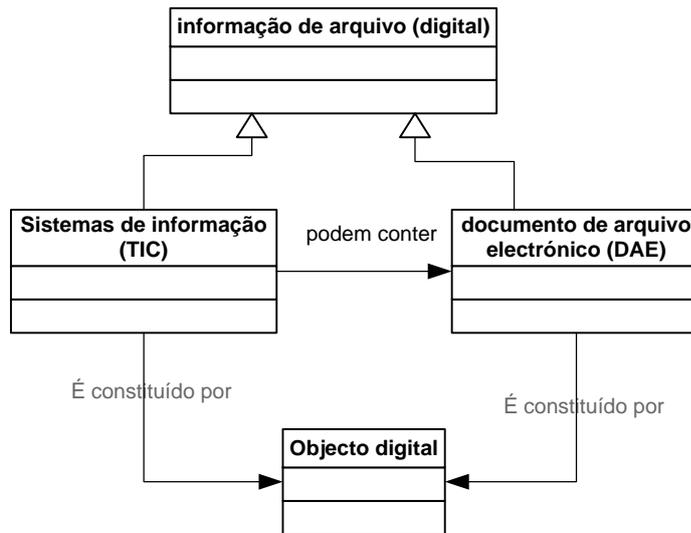
Um organismo produz informação de arquivo (iARQ): esta pode assumir as seguintes representações:

1 - Sistema de informação, mediado tecnologicamente e normalmente tributário de aplicação de metodologias e ferramentas de bases de dados. Este grupo de representações inclui sistemas que comportam, incluem ou integram em multimédia. Exemplos destes sistemas podem ser encontrados em bases de dados (de maior ou menor porte) sistemas de informação Geográfica (SIG), sistemas de webconference, etc.

2 - Documentos de arquivo eletrónicos (DAE), ou seja, objetos discretos legíveis de forma equivalente aos seus similares em suporte papel, que são, neste caso, documentos obtidos através de ferramentas de produtividade (MS Word, Excel, Powerpoint, CAD/CAM, etc.)

Os sistemas de informação podem conter ou gerir DAE. É o caso específico dos sistemas de gestão documental que integram um sistema baseado em bases de dados destinado a gerir, manipular e referenciar documentos produzidos e recebidos na organização. No contexto do presente documento apenas se irá tratar documentos de arquivo eletrónicos geridos por sistemas específicos pelo que apenas serão referidos sistemas de informação (SI)

Assim ao referirmos no presente documento a expressão “Informação de Arquivo” (iARQ) estaremos a incluir todas as suas representações desde que estas existam em formato digital.



0.1 A necessidade de um Plano de Preservação Digital (PPD)

As tecnologias de informação são, atualmente, o principal suporte para a produção e armazenamento de informação. As atividades organizacionais são, em graus de menor ou maior intensidade, mediadas tecnologicamente. Informação de diversos tipos, seja ela de apoio à decisão, operacional, probatória, ou qualquer outra, é produzida e mantida eletronicamente estando portanto dependente de um sistema intermediário composto pelo *software* e *hardware* que contribuiram para a sua criação e, naturalmente, indispensáveis para recuperar e utilizar essa informação. A rápida taxa de obsolescência tecnológica, inerente à indústria informática, levanta problemas críticos de preservação de informação operacionalmente indispensável à organização. No entanto, a tarefa de preservação digital está longe de ser simples ou isenta de custos. A experiência constatada no terreno, pela DGARQ, identificou cenários prevalentes em que a informação é produzida com carácter de utilização imediata sem serem consideradas necessidades operacionais sobre essa mesma informação a médio ou longo prazo. O resultado desta atitude resulta na perda, muitas vezes irreversível, de informação com consequências mais ou menos dramáticas para a instituição que a perdeu.

A produção do presente documento visa dar uma linha de orientação para as organizações que produzam e dependam, em maior ou menor percentagem, de informação criada e mantida eletronicamente. Essa orientação servirá para tomar medidas que possam garantir as condições materiais mínimas para preservar informação digital, durante o período pelo qual a organização dela necessite.

Não é considerada, neste documento, a circunstância em que um produtor deseje criar um arquivo digital, orientado para a preservação de objetos digitais (OD) de conservação permanente e acessível a utilizadores externos à organização. Neste caso muitos outros requisitos serão necessários bem como recursos humanos e materiais indispensáveis para gerir esse tipo de função. Neste caso, a metodologia aconselhada é a de proceder a transferências desses objetos para um arquivo digital especializado nessa função. Atualmente a DGARQ encontra-se a desenvolver um sistema desta natureza - o RODA (Repositório de Objetos Digitais Autênticos) - que já se encontra em exploração (<https://roda.dgarq.gov.pt/roda-wui>)

0.2 Estado da situação

As novas tecnologias digitais permitem novas possibilidades que, trazendo consigo benefícios e problemas, tornam os objetos delas resultantes complicados de gerir, particularmente se considerarmos períodos prolongados de tempo. Alguns exemplos:

- O mesmo documento ou formato de documento pode ser representado em diferentes suportes para diferentes finalidades, obrigando à distinção entre formato e suporte;
- Diferentes suportes podem ser utilizados em diferentes fases de criação/captura, armazenamento, acesso, distribuição e uso do documento, podendo os documentos ser codificados de maneiras completamente diferentes em cada uma destas fases;
- A separação entre formato e suporte e o tratamento separado de cada uma das fases do documento, permite uma estruturação mais livre dos documentos sem constrangimentos inerentes ao suporte papel;
- Os conteúdos originalmente digitais podem representar material de leitura, som, vídeo, fotografias, artes gráficas, radiografias, jogos, registos, dados e outros;
- A qualidade da cópia digital permanece, em tudo, idêntica, independentemente do número de cópias sucessivas realizadas;
- A transposição para papel não é aplicável por ocasionar perda de informação significativa que comprometeria decisivamente a sua utilidade e inteligibilidade¹.

O mercado informático não oferece soluções orientadas para a preservação digital, embora haja bastantes aplicações vocacionadas para gestão de documentos eletrónicos. Tais ferramentas permitem apenas gerir eficazmente funcionalidades de gestão de documentos, tais como produção, circulação ou armazenamento, não sendo previstas capacidades conducentes à preservação dos documentos ao longo de períodos de tempo prolongados.

No entanto as organizações necessitam, sob o ponto de vista operacional, da informação produzida digitalmente durante períodos mais ou menos prolongados, consoante a natureza da informação e a perenidade do seu valor administrativo, legal ou de qualquer outra natureza.

Neste cenário a preservação que, em ambiente papel, é normalmente confiada a instituições especializadas, torna-se um problema premente das organizações produtoras. É fundamental que estas sejam capazes de gerir a sua informação digital de forma a mantê-la utilizável e garantir os propósitos operacionais da organização.

O presente documento pretende contribuir para esse propósito.

¹ Exemplos típicos desta situação são por exemplo, bases de dados ou documentos multimédia

0.3 Preservação digital

A preservação digital, considerada no contexto deste documento, assume três vetores:

1. Conjunto de atividades desenvolvidas com o fim de aumentar a vida útil da informação de arquivo (iARQ), salvaguardando a utilização operacional e protegendo-os das falhas de suportes, perda física e obsolescência tecnológica;
2. Conjunto de atividades que promovem a acessibilidade continuada aos conteúdos;
3. Conjunto de atividades que assistem na preservação do conteúdo intelectual, forma, estilo, aparência e funcionalidade.

Assim sendo, para maximizar as perspetivas da preservação de OD a longo prazo, devem ser aplicadas medidas de preservação logo que possível, ou seja, preferencialmente enquanto os OD são/estão ainda acessíveis e ainda na fase de produção.

Quanto mais depressa a instituição integrar a preservação digital no seu plano de ação (e orçamento!), definir e implementar soluções apropriadas, maior a probabilidade de sucesso na preservação dos OD.

Pretende-se pois que a organização possua um documento estratégico que determine, à partida, quais os procedimentos a realizar para evitar obsolescência tecnológica e a consequente probabilidade de perder informação.

Os organismos são encorajados a serem proactivos na prossecução das suas estratégias de preservação digital e na definição e implementação dessas medidas, antes que a sua informação digital se torne obsoleta e conseqüentemente inutilizável. Na realidade as medidas de preservação são normalmente tomadas reactivamente - quando o são de todo - como resposta a situações emergentes de constatação de perda de informação necessária, quando tais medidas deveriam ser planeadas e implementadas como parte de uma solução de preservação digital a longo prazo.

0.4 O Plano de Preservação Digital

Um Plano de Preservação Digital:

1. É um documento estratégico que
2. contém políticas e procedimentos orientados
3. para a constituição de uma estrutura técnica e organizacional que
4. permita preservar de forma continuada documentos de arquivo eletrónicos (DAE) através de
5. ações realizadas sobre os objetos digitais (OD) que os compõem.

Destina-se a garantir que os DAE sejam conservados de forma legível e acessível, mantendo simultaneamente as suas propriedades de autenticidade e integridade durante tanto tempo quanto a organização deles necessitar.

O Plano de Preservação Digital permitirá identificar quais as funcionalidades que devem ser implementadas e a forma de as implementar, para manter a integridade e usabilidade dos documentos de arquivo eletrónicos ao longo do tempo.

O desenvolvimento de um Plano de Preservação Digital e a seleção das estratégias apropriadas, deve ser o resultado de um esforço de colaboração entre as unidades orgânicas referentes ao arquivo (gestão documental) e à informática (tecnologias da informação), com a participação de todas as unidades orgânicas afetadas pelo processo ou que produzam informação eletrónica.

Trata-se de um processo diretamente ligado a práticas de gestão documental, como adiante será explicado.



A preservação digital implica custos significativos que devem ser analisados e previstos pela Organização através do reforço orçamental, considerado adequado à situação diagnosticada aquando da elaboração do Plano de Preservação Digital

0.5 Objetivos e destinatários deste documento

Estas recomendações para a elaboração de planos de preservação digital destinam-se a ser utilizadas por qualquer organismo que:

- Produza documentação/informação em formato Eletrónicos (especificamente documentação de cariz arquivístico);
- Mantenha em formato Eletrónicos essa informação;
- Tenha reconhecidas necessidades de utilização operacional dessa informação que excedam sete anos de longevidade;
- Pretenda produzir um documento de planeamento que lhe permita determinar procedimentos de forma a reunir as condições materiais necessárias a preservar de facto essa informação.

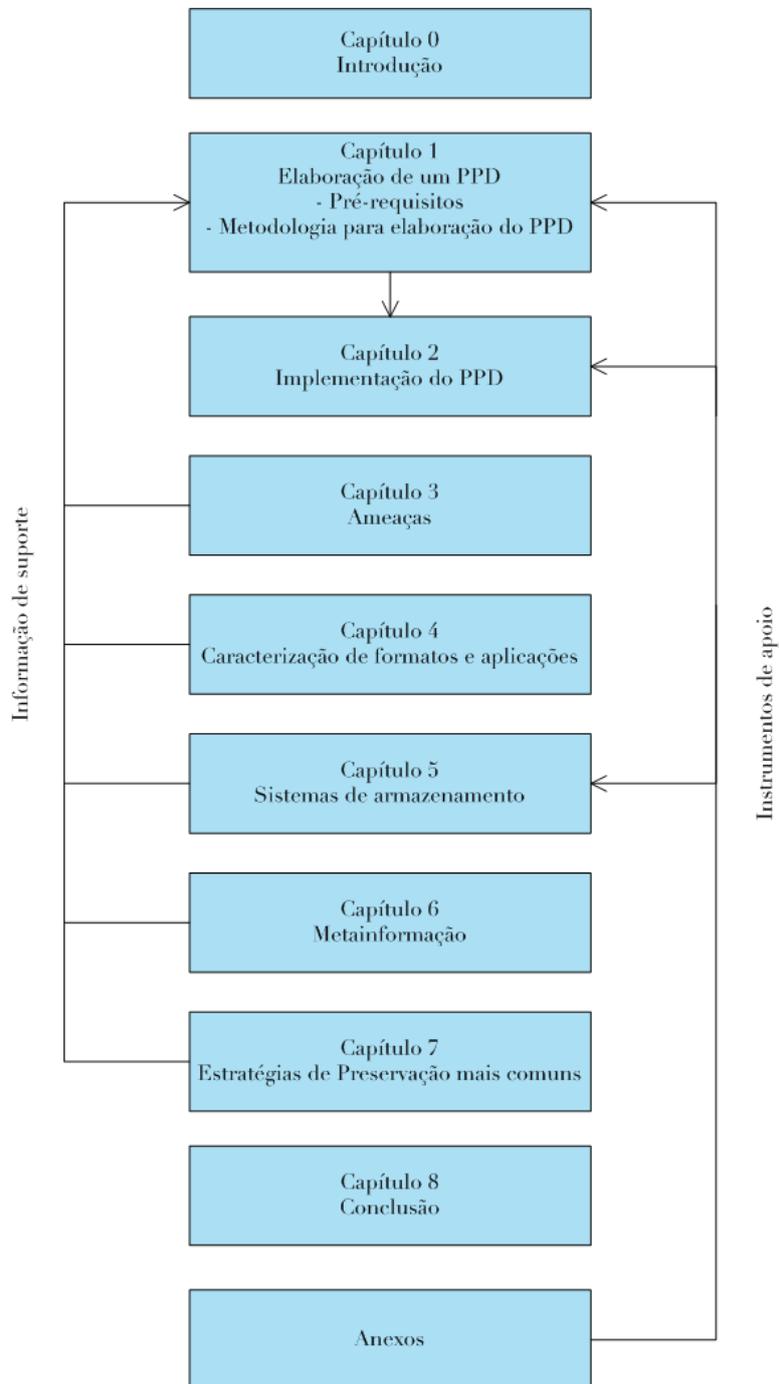
0.6 Estrutura e conteúdo do documento

Este documento é composto por nove capítulos, conforme o esquema que abaixo se apresenta. Salienta-se que, no início de cada capítulo, foi introduzido um resumo do conteúdo do mesmo.

A Introdução (Cap. 0) e a Conclusão (Cap. 8) destinam-se a contextualizar e evidenciar as problemáticas relativas a elaboração de Planos de Preservação Digital.

Entre os Cap. 1 e 2 existe uma relação cronológica sequencial, enquanto que os Cap. 3 a 7 contêm informação de suporte ao Cap. 1.

Finalmente, os Anexos contêm instrumentos de apoio aos Cap. 1, 2 e 4.



(página intencionalmente deixada em branco)

1 Elaboração de um Plano de Preservação Digital

Este capítulo encontra-se dividido em dois grandes pontos:

O primeiro destina-se a apresentar e explicar os pré-requisitos considerados como basilares para uma organização que pretenda proceder à elaboração de um PPD, nomeadamente, o plano de classificação e a tabela de seleção.

O segundo desenvolve as etapas e atividades da metodologia para a elaboração do PPD:

- A avaliação do desempenho dos sistemas de informação e dos DAE através do levantamento, análise e caracterização dos mesmos. Com vista a agilizar a recolha de dados, são fornecidos em anexo os instrumentos necessários.
- Planeamento da estratégia de preservação, a qual tem como finalidade a escolha da melhor solução de preservação para o organismo e a consequente produção do PPD. Referem-se fatores a considerar ao nível da definição de estratégias de preservação e formatos de preservação, bem como da escolha de aplicações informáticas, soluções de armazenamento e metainformação.

1.1 Pré-requisitos

Consideram-se como instrumentos estruturantes de base à elaboração de um PPD:

- O plano de classificação e
- A tabela de seleção.

A classificação destina-se a organizar a informação de acordo com critérios pré-determinados (funcionais, orgânicos, etc.) de forma a compartimentá-la para sobre ela realizar de forma eficiente algum tipo de trabalho. A existência de um sistema de classificação é essencial para a organização da informação.

Plano de classificação:

Elemento estruturante do sistema de arquivo de qualquer organismo, na medida em que se apresenta como um conjunto de regras claramente definidas que promovem a organização dos documentos de arquivo.

Fonte: Manual para a gestão de documentos, p. 2-5.

O **plano de classificação** (um tipo de sistema classificativo) permite otimizar a aplicação de um PPD, na medida em que se apresenta como uma ferramenta de gestão de documentos de arquivo que visa, nomeadamente:

- Perceber as relações entre a informação de arquivo (iARQ) gerida pelos diferentes SI;
- Assegurar a continuidade da nomenclatura dos documentos ao longo do tempo;
- Facilitar a recuperação dos documentos relativos a uma função ou a uma atividade;
- Definir os níveis de proteção e de acesso a diferentes conjuntos documentais.

Para informações adicionais sobre sistemas de classificação, consultar:

- SIADE caderno 2: Recomendações para a Gestão de Documentos de Arquivo Electrónicos (MOREQ)
- Manual para gestão de documentos
- Norma Portuguesa NP4438 - partes 1 e 2

O plano de classificação deve ser funcional, refletir as atividades da organização da qual deriva e resultar da análise das suas funções/atividades. Constitui a base crítica para a organização e recuperação dos documentos.

O sistema de **classificação funcional** tem por base as funções atribuídas a uma instituição para dar cumprimento aos fins para os quais foi criada, às atividades levadas a cabo no desempenho de cada função e às ações que as concretizam. Assim, a classificação deverá refletir as funções/atividades da organização.

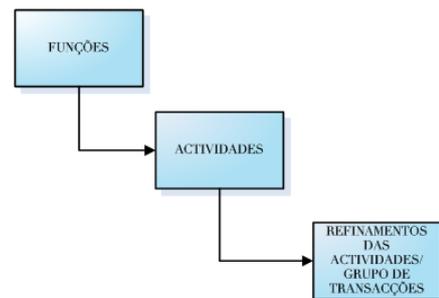
Considerando que o sistema de classificação funcional não está dependente das alterações orgânicas que ocorrem nas instituições, consegue-se uma vantajosa estabilidade organizativa dos documentos produzidos.

De acordo com a NP 4438-2², os sistemas de classificação permitem:

- Organizar, descrever e relacionar entre si diferentes grupos de informação de arquivo (iARQ);
- Relacionar e partilhar informação de arquivo de áreas diferentes, dentro da organização ou com entidades externas;
- Melhorar o acesso, a recuperação, a utilização e a disseminação da informação de arquivo

A estrutura de um sistema de classificação funcional é hierárquica, sendo que:

- O primeiro nível representa as funções,
- O segundo as atividades que constituem a função
- O terceiro e subsequentes níveis são refinamentos das atividades ou grupo de transações que decorrem no âmbito de cada atividade.



A **tabela de seleção** é uma ferramenta crucial na elaboração de um PPD, porque permite determinar os adequados prazos de retenção e ações de eliminação ou de transferência de informação de arquivo. Desta forma torna-se possível circunscrever os SI e DAE a serem abrangidos pelo PPD, i.e., aqueles cujo prazo de conservação, segundo a referida tabela, seja superior a sete anos.

Tabela de seleção:

Regista o resultado do processo de avaliação documental e fixa, para os grupos de informação identificados (séries), os respetivos prazos de conservação administrativa e destino final.

² INSTITUTO PORTUGUÊS DA QUALIDADE - NP 4438-2, *Informação e documentação. Gestão de documentos de arquivo. Parte 2: Recomendações de aplicação*. Lisboa: Instituto Português da Qualidade, 2005, p.17



Qualquer tabela de seleção que careça de aprovação da DGRQ, ao abrigo do Decreto-Lei 447/88, de 10 de Dezembro, e que inclua documentos de arquivo eletrónicos implica a elaboração de um PPD desde que se verifique uma das seguintes circunstâncias:

1. O prazo de conservação previsto seja superior a sete anos;
2. O destino final seja de conservação permanente, ou conservação permanente parcial.

Considerando que as atividades relativas à preservação de OD são dispendiosas, pois implicam a mobilização de recursos materiais e humanos especializados, assim como a afetação de tempo para as concretizar, a rigorosa aplicação da tabela de seleção assume óbvias vantagens organizativas e orçamentais.

A criticidade da informação de arquivo sob o ponto de vista probatório e a sua longevidade, (i.e., o tempo de vida operacional considerado como necessário pela organização para a sua manutenção) são variáveis independentes que não se alteram de forma diretamente proporcional.

Exemplo:

Uma fatura tem um elevado nível de criticidade ou relevância operacional que dura apenas um curto período de tempo.

Um contrato, no entanto, pode ter um grau de criticidade probatória menos elevado mas bastante mais prolongado no tempo (pelo menos durante o período de vigência da transação veiculada no contrato).

Em termos de preservação digital este aspecto adquire especial relevância. De facto se tiver sido avaliada a informação produzida na organização sob o ponto de vista de relevância operacional e longevidade necessária, podemos facilmente determinar sobre que classes de informação e sobre que sistemas de informação deveremos prever ações de preservação digital.

Exemplo 1:

Um-SI contém informação sobre execução orçamental. Essa informação de acordo com a Lei tem um prazo de conservação que atinge os dez anos, podendo a partir daí ser eliminada. Neste caso não se justifica investir tempo e recursos para preservar essa informação pois de qualquer forma poderá ser eliminada após 10 anos de vigência.

Exemplo 2

Um sistema de *webconference* através do qual se processam reuniões à distância envolvendo diversos serviços deslocalizados de uma organização. Neste contexto as gravações realizadas equivalem às *atas* e por esse motivo deverão ser objeto de conservação permanente. Neste caso é necessário prever medidas de preservação digital que permitam a manutenção desta informação.

É muito pouco provável que, no que respeita a informação de arquivo eletrónica, haja uma longevidade única, isto é, um único e idêntico período temporal durante o qual seja necessário guardar a informação. Assim, apenas interessa planejar e empreender ações de preservação sobre as classes ou sistemas de informação cuja necessidade de utilização

operacional exceda um determinado período de tempo, a partir do qual os formatos em que a informação foi produzida se começam a tornar obsoletos. Chamaremos a essa “barreira” temporal o **horizonte de obsolescência** de um determinado OD.

O horizonte de obsolescência é **estimado em sete anos**, pelo que toda a informação digital cujo prazo de vida operacional não exceda este período de tempo não deverá ser abrangida pelo PPD.

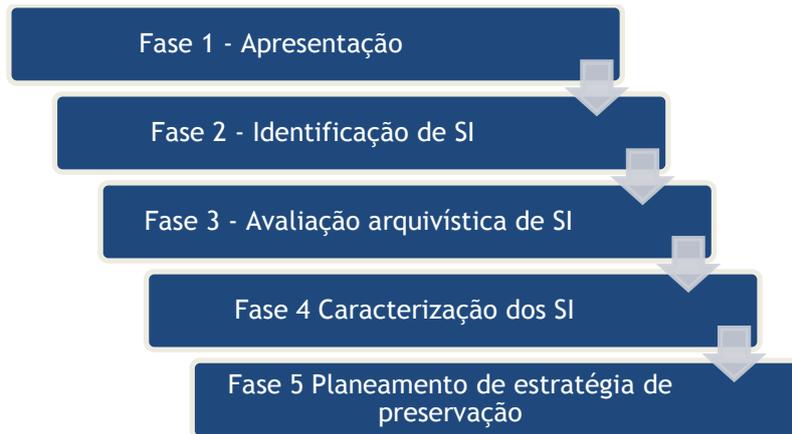
Note-se que este período não é rígido reportando-se ao prazo de Retrocompatibilidade normalmente assegurado pelas empresas vendedoras, pelo que pode ser alterado consoante a realidade específica de cada organização. Por exemplo, os sistemas de informação que suportam a atividade contabilística contêm informação cujo prazo de vida útil esperado é de 10 anos pelo que, apesar de ultrapassar o prazo sugerido, pode não ser necessário considerar estes sistemas para serem objeto de ações de preservação.

Salienta-se que os dois requisitos acima descritos são parte integrante da “Lista de requisitos para a prática de preservação digital” (Anexo A).

1.2 Metodologia para a elaboração do PPD

Apresenta-se de seguida uma tabela com as fases do processo de elaboração do processo, contemplando os atores intervenientes e todas as atividades que o compõem.

Fases do processo	Intervenientes	
	Entidade apoiada	DGARQ
1. Apresentação da metodologia para elaboração de PPD		X
2. Identificação dos sistemas de informação		
2.1. Preenchimento da FRD <u>fase 2</u>	X	
2.2. Reuniões para apoio à identificação dos sistemas de informação	X	X
3. Avaliação arquivística dos sistemas de informação		
3.1. Preenchimento da FRD <u>fase 3A (sistemas)</u> . Se necessário, preenchimento, também, da FRD <u>fase 3B (séries)</u> .	X	
3.2. Reuniões para apoio à avaliação arquivística dos sistemas (e, se necessário, das séries)	X	X
4. Caracterização e avaliação geral dos sistemas de informação		
4.1. Preenchimento da FRD fase 4A - Caract. Técnicas do Sistema	X	
4.2. Preenchimento da FRD <u>fase 4B - Síntese</u>	X	
4.3. Reuniões para apoio à caracterização dos sistemas de informação	X	X
5. Planeamento da estratégia de preservação no âmbito da elaboração do PPD		
5.1. Emissão de Recomendações		X
5.2. Elaboração da versão pré-final do PPD	X	
5.3. Validação pelos serviços	X	
5.4. Envio de projeto de PPD para Parecer	X	
5.5. Reuniões para análise técnica do projeto de PPD	X	X
5.6. Emissão de Parecer		X



1.2.1 - Fase 1: Apresentação de metodologia para elaboração de um PPD

Esta fase compreende a apresentação das fases do projeto e o papel dos seus intervenientes na produção do documento PPD. Aqui será fornecida a documentação de apoio e as instruções para a recolha da informação relativa às fases seguintes.

1.2.2 - Fase 2: Identificação dos sistemas de informação que vão ser objeto de PPD

Esta etapa tem por objetivo identificar todos os sistemas de informação existentes (Fase 1) e, através de um processo de avaliação arquivística (fase 2), selecionar aqueles que serão objeto do plano de preservação digital (por suscitarem preocupações de preservação a longo prazo - mais de 7 anos).

Informação de Arquivo: Informação registada num suporte, e que:

- produzida e/ou utilizada no âmbito das atividades da organização;
- serve de prova às atividades dessa organização

Os sistemas de informação produzem ou armazenam dados e informação, transacional ou analítica, diretamente relacionada com as atividades desempenhadas pela organização. Neste sentido o seu desempenho terá impacto nos processos a utilizar para preservar essa informação.

Informação de arquivo eletrónica - Conjunto de dados e/ou objetos digitais (p. ex. ficheiros num repositório de um SI) integrados num SI que são produzidos e/ou utilizados no âmbito das atividades da organização e servem de prova às atividades dessa organização.

Um sistema de informação adquire e gere os dados necessários para:

- A condução das atividades da organização;
- A perceção das necessidades e expectativas dos clientes internos e externos;
- Solidificar posições num ambiente competitivo;
- Apoiar as atividades dos atores organizacionais;
- Adaptar a organização à envolvente (diminuir a incerteza e a complexidade).

O desenvolvimento de um PPD implica a análise da informação de arquivo eletrónica, produzida pela organização, e que obedece ao seguinte conjunto de propriedades:

- Ser produzida na rede organizacional;
- Ser considerada como propriedade institucional;
- Ter um prazo de conservação superior a sete anos, de acordo com a respetiva tabela de seleção (horizonte de obsolescência);
- Possuir reconhecida importância (vital) de acordo com tabela de seleção, análise de risco e criticidade de sistemas para o organismo.

Isto significa que nem toda a informação de arquivo eletrónica produzida na organização deverá ser objeto de análise para a elaboração do PPD. Na organização existem, certamente, centenas ou milhares de DAE produzidos a título individual, como documentos pessoais de trabalho, que não foram comunicados e que nessa qualidade não têm representatividade institucional. Da mesma forma há, muitas vezes, pequenas aplicações, bases de dados ou sistemas de informação desenvolvidos a título pessoal pelos colaboradores para os auxiliar nas suas atividades e que podem não ter expressão institucional, ou seja, não constituem para todos os efeitos informação de arquivo.

O conhecimento do sistema é uma fase essencial no processo, pois permite aferir a situação real no que respeita à existência de condições viáveis para assegurar a preservação da informação de arquivo eletrónica, registando os pontos fortes e fracos da organização no que respeita a preservação digital e indicando quais as áreas a desenvolver para implementar o PPD.

Para tal deverá ser realizado o levantamento de todos os SISTEMAS de informação existentes no organismo, independentemente de estarem ou não em exploração (Fase 2), tendo por base a folha 1 - Identificação da Folha de Recolha de Dados que se apresenta no **anexo B**.

No quadro de elaboração do PPD e numa perspetiva operativa de delimitação do objeto a analisar, considera-se um sistema de informação um **sistema informático que, conceptual e fisicamente, se constitui como unidade autónoma ou autonomizável de outros SI**. Os SI têm, frequentemente, expressão informacional em bases de dados ou sistemas de bases de dados relacionadas, mas podem também ser constituídos por ficheiros (a informação é armazenada no filesystem). Quando o SI seja constituído por diferentes subsistemas (módulos ou mesmo aplicações específicas que processam a entrada de dados, armazenamento, recuperação, análise e exploração, reporting, ou outras funções), estes não devem ser considerados individualmente, nem para efeitos de identificação do sistema (Fase 2) nem para efeitos de avaliação arquivística (Fase 3). Contudo, relativamente à caracterização tecnológica (Fase 4A) ou, em casos excecionais para efeitos de avaliação arquivística, poder-se-á justificar a decomposição do SI.

Uma possível forma de auxiliar a este processo de delimitação é identificar dentro do sistema, a base de dados (ou seja, o repositório dos dados) que, considerando a lei de independência dos dados, se encontra separada do seu sistema de gestão: Assim uma BD tem normalmente um SGDB que sobre ela irá processar as mais variadas operações: Do ponto de vista sistémico no entanto o repositório de dados é o objeto nuclear do sistema e que servirá de pilar à delimitação do mesmo.

Deve também ser efetuada a avaliação dos sistemas de informação existentes e identificação dos que vão ser objeto de PPD (Fase 3) apenas para os sistemas que não estão referenciados em tabela de seleção em vigor, aprovada nos termos da lei (sistemas ainda não avaliados).

1.2.3 - Fase 3: Avaliação arquivística dos sistemas de informação

A fase 3 compreende as subfases:

- A - Avaliação de sistemas
- B - Avaliação de séries (apenas para as séries constantes em Sistemas de Eletrónicos de Gestão de Arquivos)

O objetivo é determinar prazos de conservação e destino final da informação que reside nos sistemas para que, posteriormente, o trabalho de elaboração do PPD incida apenas sobre os que efetivamente necessitam de ter um plano de preservação digital.

Conceito:

É necessário, ainda, explicar o que se entende, neste contexto, por série documental.

Uma série documental (independentemente do suporte) é um conjunto de documentos que obedecem a idêntica competência funcional e tipológica. Neste contexto pode tratar-se de um:

- Sistema de informação que suporte uma determinada atividade;
- Conjunto de sistemas de informação que suporte a mesma função ou atividade, ou ainda
- Conjunto de documentos eletrónicos não estruturados (por exemplo documentos de texto e mensagens de correio eletrónico que façam parte de um processo administrativo).

Alguns conceitos arquivísticos que nos são habituais e raramente suscitam dúvidas podem tornar-se confusos quando passamos do papel para o digital.

Por exemplo, um sistema de bases de dados transacional produz informação de arquivo, na medida em que os dados contidos estão vinculados a transações específicas. Qual a identificação arquivística que corresponde a esta entidade informacional? Uma série? Um documento composto? Dito de outro modo, podemos interrogar-nos sobre a equivalência de um sistema de informação relativamente a entidades conceptuais arquivísticas tradicionais.

No entanto pode-se apontar semelhanças nos dois universos. Em ambiente digital uma base de dados corresponde a um contentor onde é criada a estrutura necessária para inserir e manter dados (tabelas e relações). Na medida que assegura o suporte a uma ou mais funções, pode ser enquadrada sob uma ou mais séries. Os ficheiros que a constituem são referenciados ao nível do objeto digital, ou seja, as componentes físicas da BD. Cada registo lançado na base de dados poderá, dependendo da sua complexidade, constituir um documento simples ou composto.

Tomando como exemplo o SRH (Sistema de gestão de Recursos Humanos) teríamos:

Série: Sistema de gestão de Recursos Humanos

- ⇒ Objeto digital 1: act_admg.sql
- ⇒ Objeto digital 2: pr-m002-02288.txt
- ⇒ Documento simples 1: registo relativo ao funcionário António Silva

Para o efeito de preservação digital é essencial identificar estes três níveis de existência:

Domínio	Nível	Ação	Objetivo
Conceptual	Série	Posicionar o sistema no contexto funcional da organização e identificar, PCA prazo de conservação administrativa e destino final	Adquirir conhecimento sobre o enquadramento dos sistemas de informação existentes face à organização classificar a informação
Lógico	Sistema de informação	Delimitar o sistema de informação e caracterizá-lo funcional e tecnologicamente	Adquirir conhecimento específico e individualizado sobre os sistemas de informação existentes
Físico	Objeto digital	Identificar as componentes do sistema de informação e caracterizá-las tecnologicamente	Delinear estratégias de preservação

A observação de diversos sistemas organizacionais permite afirmar que nem sempre todas as características de um sistema de informação se aplicam a todos os OD produzidos. Efetivamente, a decisão desta escolha depende, muitas vezes, de fatores como:

- A criticidade atribuída à informação;
- A irrelevância, por vezes injustificada, atribuída a outro tipo de informação de arquivo eletrónica (por exemplo sítios *web*);
- O desconhecimento de que determinados OD são de facto produzidos na organização ou onde estes, efetivamente, se encontram.

Apenas os sistemas que, no final da fase 2, forem selecionados, serão objeto de uma análise mais detalhada na FASE 3 uma vez que se trata de uma fase diretamente dirigida à definição dos termos de preservação digital a implementar.

1.2.4 - Fase 4: Caracterização tecnológica dos sistemas de informação que vão ser objeto de PPD e sua Avaliação

A fase 4 tem por objetivo caracterizar tecnicamente os sistemas de informação selecionados na fase anterior e requer o preenchimento da FRD 4A - Caracterização Tecnológica de Sistemas, para o Sistema de Informação (globalmente considerado) OU para cada um dos Subsistemas que o integram.

A análise e tratamento dos dados desta FRD permitirá obter uma descrição das características do sistema, bem como a informação necessária sobre a informação de arquivo Eletrónicos que se pretendem preservar no âmbito do PPD.

A conjugação da caracterização técnica (do sistema globalmente considerado, ou de cada subsistema) realizada na fase 4A, com a caracterização arquivística, efetuada na fase 3, deverá permitir uma avaliação conducente à estratégia de preservação digital mais adequada para cada sistema de informação. Tal avaliação deve ser plasmada na FRD 4B - Síntese.

1.2.5 - Fase 5: Planeamento da estratégia de preservação

O planeamento da estratégia de preservação tem como finalidade a escolha da melhor solução de preservação para o organismo e a conseqüente produção do documento “Plano de Preservação Digital”, de forma a assegurar a integridade e funcionalidade continuadas da informação de arquivo eletrónica e dos objetos digitais constituintes. Essa solução comporta normalmente diferentes estratégias variáveis conforme o resultado da análise empreendida e já descrita.

O planeamento, feito com base na informação reunida, comporta:

- Os procedimentos que devem ser realizados pelo organismo sobre as diversas classes de informação identificadas (ver Cap. 7);
- A definição de formatos para preservação para cada uma das classes de informação caracterizadas (ver Cap. 4.1);
- A seriação fundamentada de soluções de software alternativas para um determinado formato (ver Cap. 4.2);
- A seleção e utilização de soluções de armazenamento adequadas a preservação digital (ver Cap. 5);
- Os esquemas de metainformação que devem ser utilizados (ver Cap. 6).

No planeamento da estratégia de preservação as instituições devem ter em conta, os seguintes fatores:

Os processos de preservação requerem a utilização de normas reconhecidas, nomeadamente de normas para gestão de documentos (Ex: NP 4438).

- Custos de implementação e custos de manutenção para as medidas de preservação;
- Complexidade técnica da abordagem escolhida e a capacidade da organização para suportar essa abordagem ao longo do tempo (técnica e financeiramente);
- Compatibilidade com o *software* e *hardware* existentes;
- Impacto nos processos de negócio (ex.: se a abordagem escolhida obriga a alterações das práticas de trabalho);

- Eficácia e robustez da abordagem na proteção da integridade, acessibilidade e funcionalidade dos documentos de arquivo eletrónicos (dos sistemas de informação?) ao longo do tempo.

Naturalmente a intensidade das medidas de preservação a empreender devem estar diretamente relacionadas com o grau de criticidade atribuída aos objetos digitais, assim como a análise de risco que deve ser efetuada.

1.2.5.1 Definição de estratégias de preservação

A definição da estratégia de preservação implica a identificação, análise, caracterização e adequação de procedimentos para a salvaguarda de informação digital, de acordo com a informação reunida e as conclusões tomadas anteriormente. Estas vão definir a especificidade da estratégia (ou conjunto de estratégias) de preservação que o organismo irá assumir e materializar no PPD.

Os elementos a identificar para a definição da estratégia de preservação estão indicados com mais detalhe no ponto 3 das *Instruções para a utilização do Anexo A - Lista de Requisitos para a prática de preservação digital*.

Adicionalmente, os custos, tanto com recursos humanos como materiais, têm uma importância elevada na escolha das soluções de preservação. É importante, nesta fase, empreender uma análise custo/benefício que compare os diversos cenários de preservação digital adequados à informação que é identificada e avaliada em função das seguintes variáveis em parte já referidas neste documento:

- Criticidade dos OD, i.e., a sua importância para o desempenho das atividades da organização;
- Longevidade operacional, ou seja, qual o tempo de vida operacional atribuído ao objeto digital;
- Risco aceitável pela organização, ou seja, a avaliação pela organização do grau de risco aceitável de perda ou obsolescência dos OD;
- Nível de formalização da informação de arquivo eletrónica produzidos (ignorar informação de arquivo produzida a título individual, por ex.).

Tal análise deve ter em conta os custos correspondentes, de forma a obter uma solução valorizada, de acordo com as variáveis referidas. Aconselha-se ainda a criação e adoção de regras e procedimentos, de forma a ajudar a normalizar a aplicação das técnicas escolhidas por todos os sistemas da instituição.

Uma abordagem mais específica das estratégias encontra-se no Cap. 7.



1.2.5.2 Definição de metodologia de manutenção da legibilidade da informação

A preservação digital só faz sentido para garantir a manutenção da acessibilidade, visualização e/ou leitura da informação. As estratégias de preservação escolhidas devem garantir que todos os dados que serão alvo de preservação estarão em condições de ser lidas e interpretadas pelos utilizadores futuros, sejam elas originárias de SI ou DAE. Para tal é necessário verificar o Prazo de Conservação Administrativa e o Destino Final da informação, no sentido de serem determinadas as especificações de preservação a nível das necessidades primárias (administrativas) e de preservação a nível das necessidades secundárias (informacionais, culturais, etc.)

As especificações têm que considerar os seguintes pontos, para cada sistema:

- 1- Estratégia;
- 2- Incidência (sobre que dados, incluindo metainformação) incide a estratégia);
- 3- Periodicidade do procedimento;
- 4- Outras observações

As especificações podem passar pela migração dos dados para outro sistema, a manutenção de sistema original, a evolução de versões (sejam dos formatos dos ficheiros ou mesmo do SI), a decretação para formatos normalizados, etc.

1.2.5.3 Definição de formatos de preservação

A escolha do formato de preservação deve ser feita antes da criação do OD que se pretende preservar evitando-se, deste modo, custos adicionais na sua conversão para o formato final. A utilização de formatos normalizados e bem documentados de aplicações com especificações abertas permitem garantir uma continuidade de utilização mais estável e prolongada, quando comparados com os formatos e as aplicações proprietários (ver Cap. 3.1).

Pretende-se que um dos resultados do PPD seja a produção de uma tabela que relacione a cada um dos formatos utilizado na Organização, um ou mais formatos de preservação definidos para o efeito.

Apresenta-se uma lista de categorias de formatos no Cap. 4.1.

1.2.5.4 Escolha de aplicações informáticas

A escolha do sistema aplicacional da organização deverá ter em conta, como um dos critérios de escolha, as vantagens que a aplicação oferece para a estabilização e normalização da informação produzida através dela.

Dependendo da informação recolhida no âmbito da FRD fase 4 - Caract. Técnicas do Sistema:

- Poderá ser necessário proceder à aquisição de *software*, caso o existente não dê resposta às medidas previstas no PPD;
- Poderá ser necessário proceder à adaptação, alteração ou substituição de *software* utilizado no Organismo, de acordo com o indicado no PPD;
- Deverão constar, no PPD, orientações que sirvam de base a futuras aquisições de *software*.

Em todos estes casos há que ter em consideração o conjunto de estratégias de preservação e os formatos definidos anteriormente, bem como a verificação da existência, no mercado, de soluções compatíveis com essas estratégias e formatos. Caso estas não existam ou não sejam satisfatórias, a opção passa por se investir no desenvolvimento de soluções específicas, ou empreender ações de “congelamento” dos OD produzidos nesse formato até ao aparecimento de uma solução viável.

Neste último caso aconselha-se a utilização, sempre que possível, de aplicações informáticas com especificações abertas, não proprietárias e bem documentadas, pelas razões indicadas no Cap. 3.1. Estas devem permitir executar todas as operações necessárias definidas no âmbito do PPD, ser robustas, fáceis de utilizar e conter uma interface intuitiva.

Identificamos as tipologias de aplicações de software quanto à sua origem no Cap. 4.2.

1.2.5.5 Escolha de soluções de armazenamento

O sistema de armazenamento utilizado é de importância crítica para a preservação digital. É com efeito imprescindível que este obedeça a características que assegurem a persistência e segurança dos objetos digitais a preservar. É ainda preciso notar que o grande volume de informação digital normalmente produzida exige elevada capacidade de armazenamento acompanhada dos imprescindíveis mecanismos de segurança (entre os quais o backup se conta). A escolha da solução de armazenamento deve considerar os recursos e as infraestruturas de apoio existentes, de forma a avaliar:

- A sua adequação a necessidades de preservação digital;
- Quais as necessidades de aquisição de sistemas de armazenamento adequados (no caso da resposta ao quesito anterior ser negativa);
- Os custos com necessidades de armazenamento de informação a médio/longo prazo.

Uma opção que se venha a revelar menos adequada, poderá ter como resultado indesejável obrigar a uma inflexão de todo o planeamento elaborado. Uma análise mais aprofundada sobre as soluções de armazenamento aconselhadas encontra-se no Cap. 5.

1.2.5.6 Escolha de Metainformação

A metainformação tem também um papel primordial na preservação digital, visto que será esta que permitirá uma melhor gestão dos:

- Objetos Digitais;
- Sistemas de informação (por exemplo registo de correspondência, Sistemas de informação geográficos);
- Utilizadores;
- Funcionalidades a disponibilizar.

Por exemplo o sistema de registo de documentos utiliza metainformação, os serviços de Tecnologias da Informação e Comunicação (TIC), utilizam-na igualmente quando guardam rotinas de auditoria que registam o comportamento dos sistemas e os acessos aos mesmos.

A metainformação é normalmente utilizada dentro da organização - seja de forma implícita ou explícita - com diversos propósitos e por diferentes áreas funcionais. Basicamente pretende-se referenciar ou descrever outra informação de forma a ser mais fácil a sua posterior recuperação ou ainda garantir a evidência de uma determinada transação (veja-se por exemplo a metainformação embebida no cabeçalho de imagens ou nos documentos de texto). Na maior parte dos casos observados não existe, todavia, uma formalização organizacional (num documento único por exemplo) de todos os esquemas de metainformação utilizados e de quais os respetivos elementos. Propõe-se assim a criação de um documento deste tipo, de forma a oficializar e disseminar a utilização de metainformação e evitar redundâncias e repetições muitas vezes verificadas.

Tal documento deve adequar-se igualmente à documentação analógica e à documentação digital. A vantagem desta prática reside em saber exatamente que metainformação é utilizada para cada classe de informação.

Os processos de preservação devem estar totalmente documentados e a documentação guardada para apoio a futuros esforços de preservação. Qualquer cópia ou reformatação de dados para migração ou conversão deve ser documentada na metainformação de gestão documental.

A utilização de metainformação específica de preservação digital é ainda uma das componentes que deve ser integrada nesse documento.

2 Implementação do Plano de Preservação Digital

Neste capítulo procede-se à abordagem da implementação do Plano de Preservação Digital, mencionando algumas recomendações a tomar aos seguintes níveis: teste da solução proposta; produção de cópias de segurança; aplicação da estratégia de preservação; medidas de monitorização e auditorias internas; e afetação de recursos.

São apresentadas, de seguida, algumas recomendações de ações a tomar na fase de implementação do PPD. Estas não pretendem ser rígidas, visto que a implementação do plano de preservação digital, ou seja, a aplicação prática no terreno das medidas nele preconizadas, deverá ser programada pela própria instituição. Aconselha-se a utilização de bibliografia ligada à temática de Gestão de Projeto³, como apoio a esta fase.



A DGARQ, como entidade coordenadora do sistema nacional de arquivos, é neutra quanto ao método escolhido pela organização no que concerne à implementação do PPD. À DGARQ importa apenas, em sede de auditoria, o resultado dessa mesma implementação. Nesse âmbito, o anexo A estabelece uma **lista de requisitos considerados obrigatórios** para a prática de preservação digital, a qual servirá de base a uma eventual auditoria.

Este facto não inviabiliza que a DGARQ forneça apoio técnico para a construção de uma estratégia de preservação digital quando uma organização o solicite.

2.1 Teste da solução proposta

Antes da implementação do PPD, os procedimentos técnicos e soluções decididas a partir do processo de avaliação descrito em 1.2.2 devem ser sujeitos a testes. Estes devem ser levados a cabo pelos colaboradores do organismo sobre duplicados dos documentos de arquivo eletrónicos de forma a salvaguardar os originais de quaisquer resultados indesejáveis, sempre possíveis em fases de teste.

2.2 Produção de cópias de segurança

Antes da intervenção de implementação devem ser feitas cópias de segurança de todos os OD selecionados para preservação. A integridade destas cópias deve ser verificada antes da sua remoção para uma área de armazenamento segura. Estes duplicados não devem ser sujeitos a qualquer tratamento de preservação, devendo funcionar como matrizes caso os procedimentos preconizados para a estratégia de salvaguarda sejam mal sucedidos.

³ Por exemplo: ANSI/PMI - A Guide to the Project Management Body of Knowledge: PMOBOK Guide, 3ª ed., Pennsylvania: Project Management Institute, 2004. ISBN 193069945X

2.3 Aplicação da Estratégia de Preservação

Após a finalização desta fase e caso os resultados tenham sido satisfatórios, o tratamento de preservação deve ser estendido a todos os OD selecionados para tal.

2.4 Medidas de monitorização e auditorias internas

A eficácia da aplicação do plano de preservação digital deverá ser regularmente monitorizada internamente considerando o conjunto de tópicos que compõem estas recomendações descritas em 1. Desta forma será possível identificar eventuais disfunções e empreender ações corretivas.

Tópicos a considerar na Monitorização:

- Os procedimentos de preservação a realizar pelo organismo (ver Cap. 7);
- A definição de formatos para preservação; (ver ponto 3.1)
- A listagem de soluções de software alternativas para um determinado formato (ver ponto 3.2);
- A seleção e utilização de soluções de armazenamento; (ver ponto 5)
- Esquemas de metainformação a utilizar (ver ponto 6)

Depois de sujeitos a ações de preservação, os OD deverão ser reavaliados de forma a confirmar a permanência das suas funcionalidades de acordo com os padrões pré-definidos.

A integridade de toda a metainformação relevante associada aos OD deve ser verificada após a aplicação da estratégia de preservação. A metainformação deve também ser atualizada, registando as ocorrências verificadas durante o processo de preservação de forma a constituir um histórico essencial para provar a integridade do sistema ou a verificação e comprovação de eventuais disfunções.

O que monitorizar nos documentos de arquivo eletrónicos preservados:

- Integridade;
- Funcionalidade;
- Estrutura;
- Conteúdo;
- Metainformação associada

Por exemplo.

- ⇒ Após a migração de um sistema de bases de dados produzido originalmente em ORACLE, para um formato em MySQL ou DBML, é necessário verificar se os dados migrados mantêm a sua qualidade e se as relações existentes entre tabelas permanecem explícitas.
- ⇒ Após a migração de documentos produzidos em Word para formato PDF é necessário verificar se não houve desformatação, se as notas de pé de página não foram apagadas, se os cabeçalhos permanecem, etc.
- ⇒ Após a migração de documentos de áudio e vídeo para formato MPEG é necessário verificar se não houve dessincronização entre as imagens e o som com, se não existem fenómenos de arrastamento de *pixels*, etc.
- ⇒ Após migração de publicações Multimédia Interativas é necessário verificar se as hiperligações se mantêm corretas, se é possível o acesso ao conteúdo dos ficheiros, se todos os ficheiros permanecem presentes, etc.

2.5. Afetação de recursos

É essencial prever antecipadamente os recursos humanos, materiais, bem como os elementos de gestão basilares (cronograma), considerados como necessários para assegurar o esforço de implementação do plano. A organização deve planear e implementar as estruturas necessárias para aplicar a monitorizar a aplicação do PPD.

3 Ameaças

Este capítulo destina-se a explicitar o conjunto de ameaças que pesam sobre os OD, abordando a problemática da obsolescência dos formatos, software, hardware e suportes, assim como elementos que afetam os suportes.

A especificidade dos OD implica que estes estejam sujeitos às seguintes ameaças:

■ Obsolescência tecnológica:

Ocorre ao nível de formatos e *software*, mas também do próprio *hardware* e suportes utilizados.

■ Danos físicos:

Estes podem ser acidentais ou propositados, ocorrem principalmente ao nível dos materiais que compõem o *hardware* e os suportes, mas também podem ocorrer nos ficheiros (muitas vezes como consequência dos danos de *hardware* e/ou dos suportes).

3.1 Obsolescência de formatos e software

A obsolescência de formatos e *software* pode ocorrer quando:

- Um formato de ficheiro é ultrapassado por outro formato ou por uma versão mais recente que comporta mais complexidade;
- Um dado formato não vinga ou as empresas não criam *software* compatível;
- Um dado formato falha, estagna, ou já não é compatível com os sistemas atuais;
- O *software* que suporta o formato falha comercialmente ou é adquirido por um concorrente que o retira do mercado.

No que diz respeito ao *software*, este pode dividir-se em três grupos, de acordo com o exposto no Cap. 4.2:

- Proprietário com especificações fechadas;
- Proprietário com especificações abertas;
- Não proprietário com especificações abertas.

Para escolher o melhor *software* no âmbito do planeamento da estratégia de preservação é necessário avaliar se este:

- Tem grande adoção;
- É retro compatível;
- Tem suporte de metainformação (como XML);
- Dispõe de várias funcionalidades, sem ser muito complexo;

- Permite a utilização de vários formatos;
- Inclui funcionalidades de verificação de erros;
- Dispõe de um ciclo de atualizações razoável.

3.2 *Obsolescência do hardware e suportes*

No âmbito do *hardware* e dos suportes, a rápida obsolescência é uma constante, devido aos avanços tecnológicos a nível de:

- Velocidade do processador;
- Densidade dos *chips* de memória;
- Capacidade dos periféricos de armazenamento;
- Velocidade de processamento de imagem;
- Velocidade de transmissão.

Tais alterações levam à rápida substituição dos suportes, não só devido à maior rapidez, capacidade e produtividade disponíveis, mas também devido ao cada vez maior número de funções oferecidas.

As normas ISO/IEC 17799, ISO/IEC 27000 a ISO/IEC 27008 fornecem informações, códigos de práticas, e requisitos para auditoria e certificação de técnicas de segurança para sistemas de gestão de informação. De referir, no entanto, que a análise destas normas não está no âmbito deste documento e que não cabe à DGARQ a certificação destas técnicas.

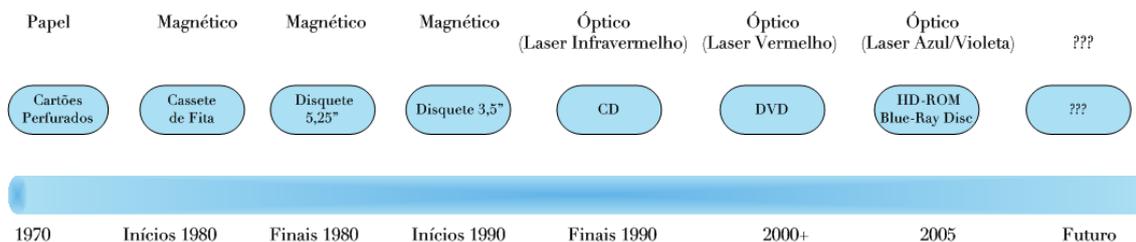
Elementos como:

- A diminuição do tamanho físico;
- O aumento da capacidade e diminuição do custo por unidade de armazenamento;
- As tendências em termos de fragilidade, estabilidade, segurança e duração antes da obsolescência e
- Outras melhorias,

Levam a novas e melhores aplicações de *software*, que aumentam a probabilidade obsolescência do equipamento anterior.

Assim os dispositivos necessários para leitura do suporte de armazenamento podem deixar de ser produzidos. Os fabricantes de todas estas tecnologias competem, emergem, unem-se, ou desaparecem, tornando cada vez mais difícil a manutenção dos conteúdos digitais ao longo do tempo.

Cronologia: evolução de suportes



3.3 Elementos que afetam os suportes

Os principais elementos que afetam os suportes e, diretamente, os OD que se pretendem preservar, são:

- A instabilidade do material;
- As condições ambientais de armazenamento (temperatura, humidade relativa e exposição à luz);
- O manuseamento e utilização (riscos em suportes óticos, dedadas, borrões, sujidade, poeiras, água e humidade, solventes orgânicos, marcadores de escrita, etiquetas adesivas);
- Desastres naturais (fogo, inundações, abalos de terra);
- Falhas de infraestrutura (canalização, eletricidade, sistema de controlo climatérico);
- Manutenção inadequada e mau funcionamento do *hardware*;
- Erro humano, sabotagem (roubo e vandalismo);
- Campos magnéticos, raios - X e micro-ondas;
- Desgaste derivado da utilização.

4 Caracterização de formatos e aplicações

Este capítulo pretende caracterizar o conjunto de formatos e aplicações existentes, sendo definidos:

- Os formatos, as suas características e propostas para estratégias de preservação;
- As diferentes categorias de aplicações e suas especificações.

São igualmente apresentados conselhos relativos formatos e aplicações a utilizar.

4.1 Formatos

Qualquer tentativa de categorização/compartimentação de tipos de documentos eletrónicos é necessariamente artificial, devido ao facto de muitos conterem, em si, elementos de outras categorias. Por exemplo, são raros os documentos eletrónicos de texto estruturado que não incluam atualmente tabelas ou imagens; são raros os formatos de registos de vídeo que não incluam também som e/ou voz. Visto que, para efeitos de definição de formatos, é necessário definir tais categorias, optou-se pela taxionomia dos tipos de documentos de arquivo eletrónicos⁴ baseada na definida pelos Arquivos Nacionais do Reino Unido⁵.

Na tabela abaixo apresentada, inclui-se uma coluna relativa à categoria de formato utilizada, uma segunda coluna que descreve as principais características dessa categoria e uma terceira coluna onde se indica a estratégia de preservação recomendada. A última coluna apresenta alguma informação adicional.

Categoria	Características	Estratégia de Preservação Recomendada	Informação Adicional
1. Dados Tabulares (<i>Data Set</i>)	<p>Dados alfanuméricos existentes numa vasta variedade de aplicações de processamento de dados;</p> <p>Dados geridos em ficheiro linear (<i>flatfile</i>), em rede, hierárquica;</p> <p>Bases de dados relacionais e orientadas a objetos.</p>	<p>Migração para novas plataformas.</p> <p>Conversão de dados e estrutura para formatos normalizados.</p> <p>Atualmente é impossível migrar as componentes comportamentais relacionadas com motores de gestão de bases de dados, normalmente proprietários.</p>	<p>O Arquivo Nacional dos Países Baixos produziu uma proposta para preservação de:</p> <p>Folhas de cálculo <www.digitaleduurzaamheid.nl/bibliotheek/ocs/volatility-permanence-spreadsh-en.pdf>;</p> <p>Bases de Dados <www.digitaleduurzaamheid.nl/bibliotheek/ocs/volatility-permanence-databases-en.pdf>.</p>

⁴ In STEPHENS, David O. - *Digital Preservation in the United Kingdom* in THE INFORMATION MANAGEMENT, vol. 34, nº. 4, outubro 2000

⁵ *The National Archives*, <http://www.nationalarchives.gov.uk/>

Categoria	Características	Estratégia de Preservação Recomendada	Informação Adicional
2. Texto estruturado/ documentos do Office	<p>Dados alfanuméricos;</p> <p>Dados de marcação (<i>markup</i>);</p> <p>Etiquetas para outros tipos de dados (imagens vetoriais e mapas de bits [<i>raster</i>]) existentes em processadores de texto/programas para escritório e digitalização de documentos/sistemas aplicativos de gestão de documentos eletrónicos, bases de dados relacionais, e outros ambientes de tipos de documentos de aplicações específicas.</p>	<p>Migração para novas plataformas.</p> <p>Conversão para formatos normalizados.</p> <p>Refrescamento de suportes.</p>	<p>O Arquivo Nacional dos Países Baixos produziu propostas para preservação de:</p> <p>Documentos de texto <www.digitaleduurzaa.mheid.nl/bibliotheek/docs/volatility-permanence-textdocs-en.pdf>;</p> <p>Mensagens de Correio Eletrónicos <www.digitaleduurzaa.mheid.nl/bibliotheek/docs/volatility-permanence-email-en.pdf>.</p>
3. Dados de Desenho [design data]	<p>Imagens vetoriais e de mapa de bits e dados alfanuméricos em sistemas de CAD e conversão orientada a objetos para bases de dados normalizadas.</p>	<p>Retrocompatibilidade, interoperabilidade, e formatos com emulação de tecnologia como estratégia a curto prazo para dados do produto (<i>product data</i>) em sistemas obsoletos.</p> <p>Refrescamento de suportes.</p>	
4. Apresentações	<p>Dados alfanuméricos em imagens empresariais, <i>clipart</i> [coleção de imagens], vídeo, e multimédia de treino/ensino.</p> <p>Diagramas para a gestão de informações, de conhecimento e de capital intelectual; para a compreensão e solução de problemas.</p>	<p>Retrocompatibilidade.</p> <p>Conversão para formatos normalizados.</p> <p>Refrescamento de suportes.</p>	
5. Imagens	<p>Imagens de mapas de bit e dados alfanuméricos existentes no software de captura/edição de imagens, bases de dados orientadas a</p>	<p>Retrocompatibilidade.</p> <p>Conversão para formatos normalizados.</p> <p>Refrescamento de suportes.</p>	

Categoria	Características	Estratégia de Preservação Recomendada	Informação Adicional
	objetos, relacionais ou simples de livrarias de fotos, coleções de digitais de belas-artes, coleções de imagens médicas.		
6. Documentos de som e voz	Dados de som em processamento de voz, bases de dados relacionais ou simples (<i>flat</i>) de registos áudio em coleções de música e repositórios similares.	Retrocompatibilidade. Conversão para formatos normalizados. Refreshamento de suportes (Esta é uma área especializada que necessita de trabalho adicional por especialistas de preservação digital).	
7. Documentos de Vídeo	Vídeo Digital, ecrã total (<i>fullscreen</i>)/ vídeo (em movimento) em fotogramas de vídeo digital armazenadas em ficheiros de mapas de bit e som armazenadas em bases de dados relacionais ou simples (<i>flat</i>) em coleções de música e repositórios similares	Retrocompatibilidade. Conversão para formatos normalizados. Refreshamento de suportes. A tecnologia de preservação é necessária em situações a curto prazo quando os dados estão fechados em sistemas proprietários.	
8. Dados geográfico/cartográfico	Imagens vetoriais e de mapa de bits e dados alfanuméricos em sistemas de informação geográfica (SIG) e <i>software</i> de cartografia. Dados armazenados orientada em bases de dados relacionais orientadas a objetos.	Retrocompatibilidade. Conversão para formatos normalizados. Refreshamento de suportes. A tecnologia de preservação é necessária em situações a curto prazo quando os dados estão fechados em sistemas proprietários.	
9. Publicações Multimédia Interativas	Dados de som e vídeo, imagens em movimento, imagens vetoriais e de mapa de bits e dados alfanuméricos armazenados em <i>software</i> de criação e edição em ambientes de publicação eletrónica.	A tecnologia de preservação é necessária em situações a curto prazo quando os dados estão fechados em sistemas proprietários até que existam normas aceites e amplamente utilizadas.	O Arquivo Nacional dos Países Baixos aborda a problemática da preservação de sítios Web em < www.digitaleduurzaamheid.nl/index.cfm?pagina=374 >.

Categoria	Características	Estratégia de Preservação Recomendada	Informação Adicional
10. Dados Científicos e Financeiros	Dados específicos de carácter científico e financeiro.	A tecnologia de preservação é necessária em situações a curto prazo quando os dados estão fechados em sistemas proprietários até que existam normas aceites e amplamente utilizadas.	

Apresentam-se, a seguir, dois exemplos que ilustram relações possíveis entre categorias que podem coexistir num mesmo formato.

Importa referir que qualquer esquema que tente representar relações entre categorias não está, nem pode estar nunca, encerrado, uma vez que a listagem de categorias existentes e as suas características distintivas nunca estarão definitivamente estabilizadas. As necessidades nesta área nunca estarão, por definição, satisfeitas assim como as respostas a essas necessidades podem ser imprevisíveis em termos de possibilidades de combinação.

Assim é natural que, no futuro, surjam novos formatos que, à partida, não sejam passíveis de categorizar na lista atual. Tal remete para a perenidade deste documento e para a necessidade constante da sua atualização.

■ Exemplo 1: Documento de texto

Dados Estruturados/Documentos de Office: um documento de texto pode conter, para além do texto, Dados Tabulares (tabelas), Imagens e ainda Dados de Desenho.



Exemplo 2: *Web-conference*

Dentro das Publicações Multimédia Interativas podemos incluir a *Web-conference*, a qual pode conter Registos de som e voz, Registos de vídeo, Apresentações, Imagens, Texto Estruturado.



4.1.1 Identificação de formatos alternativos normalizados

A criação de documentos de arquivo eletrónicos e SI, de acordo com as necessidades operativas do Organismo, obriga à utilização de uma panóplia de formatos decorrentes das aplicações de *software* utilizado na sua produção. Muitas vezes isso traduz-se na produção de formatos proprietários e não normalizados, o que dificulta a sua preservação a longo prazo.

No caso das categorias texto estruturado/documentos do Office e imagens a escolha é facilitada pela existência de formatos normalizados muito utilizados e que mantêm o conteúdo informacional e a estrutura dos documentos (mais concretamente o PDF e o TIFF). Outras categorias, como *Data Set*, *Design Data*, Dados geográficos/cartográficos e Publicações Multimédia Interativas, ainda não dispõem de formatos normalizados que garantam a preservação de todas as características dos OD.

Nestes casos é fundamental prever quais os formatos normalizados a utilizar no âmbito da preservação digital.

Recomenda-se a não utilização de formatos proprietários para armazenamento e preservação digital a longo prazo, pelas razões indicadas nos Cap. 3.1, devendo ser necessário identificar, à partida, quais os formatos que, por serem normalizados e/ou não proprietários e de fácil

preservação a longo prazo, se poderão utilizar alternativamente. Tal escolha deve assentar em primeiro lugar na categoria de formatos em que o DAE se insere.

É preciso salientar que não existem formatos normalizados para cada uma das categorias de formatos identificados. Tal comporta riscos para a preservação desses OD que não podem ser evitados. No caso em que não haja formato alternativo e o DAE/SI se encontre inactivo há que manter o formato original sujeito a monitorização regular até ser desenvolvido um formato normalizado aplicável ao formato original em questão.

Deve ser produzida uma tabela que permita definir quais os formatos de preservação que o Organismo pretende utilizar.

Reproduzimos agora uma listagem de critérios de apoio/suporte à avaliação de formatos⁶

O projeto RODA definiu para as taxonomias de OD com que opera atualmente, os seguintes formatos de preservação:

- Texto estruturado: PDF/A (www.pdfa.org)
- Imagens fixas bidimensionais: TIFF (<http://partners.adobe.com/public/developer/tiff>)
- Bases de dados relacionais: DBML, um formato XML (<http://hdl.handle.net/1822/601>)

Nome do critério	Descrição
Quota de Mercado	Se o formato é amplamente aceite ou simplesmente um formato de nicho. A quota de mercado também é conhecida como grau de “adoção”. A “adoção” refere-se ao grau de utilização do formato por parte dos criadores primários, disseminadores, e/ou utilizadores dos recursos de informação. Um elevado nível de adoção é considerado favorável para fins de preservação.
Nível de Suporte técnico	O nível de suporte técnico dado pelo criador oficial do formato. Um elevado nível de suporte é preferível num contexto de preservação.
É Normalizado (Standard)	Se o formato foi publicado por uma organização oficial de normalização. Formatos normalizados são preferíveis aos não normalizados.
Especificação Aberta	Se a especificação do formato pode ser inspecionada/verificada de forma independente. O uso de formatos abertos é fortemente recomendado em contextos de preservação.
Suporta Compressão	Se o formato suporta qualquer tipo de compressão. Formatos não comprimidos são geralmente defendidos pela comunidade.
Apenas suporta Compressão com Perda	Se o formato suporta exclusivamente um tipo de compressão que provoca perda de informação ou deterioração do objeto original. Os esquemas de compressão com perda são grandemente desaconselhados.
Suporta Transparência	Se o formato oferece funcionalidades de transparência. Este critério é relativamente específico de certo tipo de formatos (p. ex. imagens de mapa de bits). Se o formato de origem contém funcionalidades de transparência, o formato de destino deve ter também suporte para essa propriedade.
Metainformação Embebida	Se o formato contém metainformação embebida. O formato de destino deve ter capacidade de incluir/acomodar a metainformação embebida do formato de partida.
<i>Royalties</i> (taxas de utilização)	Se a utilização ou produção do formato requer o pagamento de <i>royalties</i> ou taxas de utilização. Existe preferência por formatos livres de <i>Royalties</i> .
Código Aberto	Se existem aplicações cujo código pode ser

⁶ FERREIRA, Miguel, BAPTISTA, A. A. e RAMALHO, J. C. - "An intelligent decision support system for digital preservation," International Journal on Digital Libraries, vol. 6, no. 4, pp. 295-304, 2007

Nome do critério	Descrição
	inspecionado/verificado de forma independente. A existência de aplicações de código aberto é grandemente recomendada.
Retrocompatibilidade	Se as revisões aos formatos incluem suporte para as versões anteriores. A Retrocompatibilidade é uma característica desejável.
Nível de Documentação	Se as especificações do formato estão bem documentadas. Favorece-se a existência de formatos bem documentados.
Formatos Concorrentes	Se existem formatos concorrentes ou similares. A existência de formatos concorrentes torna um formato mais atrativo para preservação, uma vez que a informação poderá ser mais facilmente convertida.
Suporte a Gestão de Direitos Digitais	Se é possível a utilização de Gestão de Direitos Digitais (DRM), encriptação ou assinaturas digitais. Desaconselha-se a existência de qualquer tipo de funcionalidade que possa constituir obstáculo no acesso à informação.
Frequência de Atualização	Qual a frequência de revisão de um formato desde a sua publicação inicial. Este critério é definido de acordo com a seguinte fórmula: número de revisões / (ano atual - ano de disponibilização). Os formatos estáveis são preferenciais. Se a frequência de revisões é muito grande, o arquivo poderá ter dificuldade em acompanhar o ritmo das mesmas.
Suporte para extensões	Se o formato permite a inclusão de extensões, tais como secções executáveis ou características marginalmente suportadas. Desaconselha-se a utilização de formatos que suportam tais funcionalidades.
Longevidade	Quantos anos passaram desde que o formato foi disponibilizado oficialmente. Os formatos de longa duração têm geralmente preferência sobre formatos novos e pouco estabelecidos.
Interpretação/descodificação transparente	Complexidade inerente à codificação: legibilidade por parte de um ser humano recorrendo a um editor do texto simples. Têm preferência os formatos que podem ser facilmente inspecionados e/ou interpretados.
Vários produtores de aplicações de leitura	Se existem várias entidades que produzem leitores/visualizadores. Para finalidades de preservação, não se deve apostar apenas em leitores produzidos somente por uma única entidade.
Várias aplicações de leitura	Se o formato pode ser lido/interpretado por diversas aplicações informáticas. Para finalidades da preservação, não se deve apostar em formatos apenas podem ser lidos/visualizados por uma aplicação específica.
Aplicações de leitura em código aberto	Se o código fonte da aplicação de leitura pode ser inspecionada/verificada de forma independente. A existência de leitores/visualizadores em código aberto é uma característica altamente desejável.
Leitor/Visualizador Multiplataforma	Se a aplicação de leitura/visualização pode ser executada, ou tem versões para várias outras plataformas (P. ex. sistemas operativos ou hardware). A existência de aplicações executáveis em plataformas concorrentes é uma característica altamente desejável num contexto de preservação.

4.2 Aplicações

As aplicações informáticas ou *software* são, conceptualmente, uma sequência de instruções a serem seguidas e/ou executadas, na manipulação, redirecionamento ou modificação de dados ou informação, e que permitem ao utilizador fazer uma ou mais tarefas específicas. Do ponto de vista técnico, integra igualmente, não só o programa de computador propriamente dito, mas também manuais, especificações, planos de teste, etc.

As aplicações fazem parte do sistema intermediário que permite a produção, manipulação, gestão e preservação dos OD.

As aplicações que produzem os documentos podem dividir-se em três tipos:

■ Proprietário com especificações fechadas:

A cópia, redistribuição ou modificação são proibidas pelo seu criador ou distribuidor. Tais restrições são conseguidas através de meios legais e/ou técnicos. Os meios legais podem incluir o licenciamento de *software*, *copyright* (direitos de autor), e uso de patentes, enquanto os meios técnicos são caracterizados pela venda ao utilizador de binários executáveis legíveis pelo *hardware*, ocultando ou encriptando o código fonte legível pelos humanos.

■ Proprietário com especificações abertas:

Vários produtores de *software* definem e utilizam especificações ou padrões para garantir a interoperabilidade entre aplicações e plataformas. Estes padrões devem estar disponíveis para livre acesso e implementação, sem quaisquer *royalties* e outras taxas e sem discriminação de uso.

■ Não proprietário com especificações abertas:

Software livre, segundo a definição criada pela Free Software Foundation, é qualquer programa de computador que pode ser usado, copiado, estudado, modificado e redistribuído sem nenhuma restrição. Este tipo de *software* opõe-se ao conceito de *software* proprietário, mas não à comercialização (*software* comercial). A forma usual de distribuição de *software* livre é anexar a este uma licença de *software* livre, e tornar o código fonte do programa disponível. Para tal é muito usual a utilização de especificações abertas (normas) para manter a interoperabilidade de todas as distribuições originárias do mesmo código fonte base.

O *software* livre permite as seguintes modalidades:

- **Gratuito (*freeware*):** não implica no pagamento de licenças de uso ou *royalties*;
- ***Shareware*:** pode ser copiado, distribuído e utilizado experimentalmente por um determinado período, após o qual se deve pagar uma taxa.

Aconselha-se a utilização, sempre que possível, de aplicações e formatos com especificações abertas, não proprietários e bem documentados, pelas razões indicadas no Cap. 3.1.

Esta prática é objeto da Resolução de Conselho de Ministros 21/2002, de 31 de Janeiro.

Para mais informações sobre a promoção da utilização de *software* livre na Administração Pública, aceda a:

<http://www.softwarelivre.citiap.gov.pt/>

5 Sistemas de armazenamento

Este capítulo destina-se a abordar os sistemas de armazenamento de acordo com um conjunto de critérios a que devem obedecer. Em seguida é feita uma análise comparativa das arquiteturas de armazenamento (NAS, SAN e CAS), sendo definidos critérios de avaliação das mesmas. Finalmente é dada relevância à problemática das cópias de segurança no que concerne à criação de réplicas, modelos de armazenamento e tipos de gestão de repositórios de dados.

5.1 Atributos desejáveis

A solução de armazenamento deve ser distinta das réplicas de segurança (também designados de ficheiros de salvaguarda ou *backups*), abordadas no Cap.5.3. Assim, um sistema de armazenamento consiste no dispositivo ou conjunto de dispositivos utilizados para armazenar primariamente os dados e a informação produzida (dispositivos primários).

O *backup* destina-se a salvaguardar os dados contidos nesses dispositivos primários. Consiste portanto numa réplica desses dados, a qual poderá ser efetuada seguindo diferentes métodos oferecendo a garantia de, em caso de perda ou corrupção dos dados primários contidos no sistema de armazenamento, estes poderem ser recuperados através das cópias de *backup* e o sistema reposto na situação original antes da rutura verificada.

Um sistema de armazenamento é uma estrutura de complexidade variável, mas não se pode confundir com um suporte de armazenamento individualizado. Por exemplo um DVD, ou um conjunto de DVD considerados individualmente sem um sistema de organização mecânica ou eletrónica, não constitui um sistema de armazenamento na aceção deste documento, embora contenha informação. Mais à frente abordar-se-á os suportes considerados de forma descontextualizada do sistema de armazenamento, pelo facto de, em muitas organizações, ainda se recorrer a esse processo para armazenar informação.

É normal um sistema de armazenamento incluir um sistema de cópia de segurança. Deverá incluir réplicas, geográfica e administrativamente independentes. Desta forma a solução de armazenamento deverá ter as seguintes características:

- Estar disponível em linha;
- Apresentar baixa latência.

Latência (em inglês *lag*): refere-se aos atrasos que se podem verificar na comunicação entre computadores (Internet, por exemplo) e comunicações via satélite. Em computação, *lag* refere-se ao tempo que um pacote de dados leva a completar um percurso de ida e volta entre um computador local e o seu destino.

A combinação destes atributos permite auditorias fáceis e frequentes e, ainda, reparações rápidas. Deve suportar a introdução progressiva de novos componentes heterogéneos (ex. discos rígidos de diferentes marcas), permitindo assim a escalabilidade e a existência de um plano de sucessão facilmente implementável.

O sistema de armazenamento deve:

- Ter um desempenho adequado às necessidades específicas de cada organização, de forma a evitar “engarramentos” (bloqueios e atrasos de reposta);
- Estar em conformidade com normas *de jure*;
- Ser independente da plataforma aplicacional utilizada (sistema operativo e vendedor de *hardware*).

Preferencialmente o custo inicial e de manutenção deverá ser mínimo, tal como o esforço necessário para a sua administração.

5.2 Arquiteturas Possíveis

As Arquiteturas de armazenamento a longo prazo, atualmente disponíveis são as seguintes:

- Network Attached Storage (NAS)
- Storage Area Network (SAN)
- Content-Addressable Storage (CAS)
- Híbrida;

Estes cenários aplicam-se a organizações com uma forte produção de DAE que requerem, portanto, elevada capacidade de armazenamento e de resposta. No caso da organização produzir DAE em pequena escala é possível recorrer a sistemas de armazenamento com desempenho inferior que podem, inclusivamente, incluir suportes óticos e dispositivos de armazenamento conexos.

Por questões que se prendem com a gestão e segurança, aconselha-se a utilização de um sistema de armazenamento, caso o volume de documentos de arquivo eletrónicos utilizados pelo organismo for superior a 10Gb.

As **NAS** são dispositivos dedicados à partilha de ficheiros entre diferentes clientes de rede. Permitem adicionar armazenamento na rede sem ser necessário desligar o servidor. Podem estar ligadas em qualquer parte da rede, não integrando, porém, o servidor que faz a gestão do processamento dos dados.

É uma solução barata, fácil de instalar, administrar e escalar, e baseia-se em protocolos abertos e bastante disseminados. No entanto a ligação entre servidor aplicacional e armazenamento é lenta; apresenta limitações no que respeita ao tamanho máximo de um bastidor (atualmente 16 terabytes); precisa de utilizar sistemas de virtualização de armazenamento (*storage virtualization*) para suportar múltiplos sistemas de ficheiros com o objetivo de guardar os dados.

As **SAN** são redes projetadas para agrupar dispositivos de armazenamentos de computador. Existem duas variações de SAN:

1. Uma rede na qual o propósito principal é a transferência de dados entre computadores e dispositivos de armazenamento. Neste caso a SAN consiste numa infraestrutura de comunicação que dota conexões físicas com uma camada de gestão, com o objetivo de organizar dispositivos de armazenamento, conexões e computadores, tornando a transferência de dados mais robusta e segura;
2. Um sistema de armazenamento formado por dispositivos de armazenamento, computadores e/ou aplicações, e todo um controlo via *software*, comunicando-se através de uma rede de computadores.

As SAN caracterizam-se por:

- Apresentar um bom desempenho;
- Ser normalmente mais dispendiosas;

- Ter maior exigência de administração (gestão de unidades lógicas e de sistemas de ficheiros);
- Ter reduzida interoperabilidade (o que funciona numa SAN pode não funcionar noutra);
- O suporte aos servidores estar dependente do sistema operativo (com as consequentes problemas de limitações de drivers de controlos).

As **CAS** são mecanismos para armazenamento de informação, cuja recuperação se baseia no seu conteúdo e não na sua localização de armazenamento. As CAS caracterizam-se por disponibilizarem um acesso simples, sem as limitações normais de um sistema de ficheiros ou protocolo de partilha.

São fáceis de instalar e administrar e têm um elevado grau de escalabilidade. No entanto são dispendiosas, pelo menos na fase inicial de aquisição, e a sua interface de programação (API) é restrita e pouco disseminada.

As soluções **Híbridas** caracterizam-se pela utilização de várias Arquiteturas de armazenamento em simultâneo e ocorrem quando se verifica grande complexidade do sistema. Tenta agregar os diversos benefícios das várias Arquiteturas utilizadas nos sistemas de armazenamento.

Apresenta-se, abaixo, um quadro comparativo destes sistemas de armazenamento:

Quadro comparativo de sistemas de armazenamento

	NAS	SAN	CAS
Heterogeneidade	Boa	Má	Nenhuma
Escalabilidade	Adicionando mais NAS	<i>Upgrade</i>	Até 24 Petabytes
Desempenho	Baixo	Muito Alto	Alto
Segue normas <i>standard</i>	Sim	<i>"Flavours"</i> da norma <i>standard</i> de base	API restrita
Independência	Completa	Parcial. Só alguns sistemas operativos são suportados	Sim. Desde que implemente a API
Custo	Baixo	Alto	Muito Alto (inicial)
Administração	Depende da virtualização	Difícil	Muito Fácil

5.2.1 Critérios para avaliação de soluções de armazenamento

São seguidamente referidas as variáveis que deverão ser consideradas para avaliar a solução de armazenamento mais adequada às necessidades de preservação identificadas. Sem detrimento de outros critérios que a organização pretenda considerar os que são a seguir enunciados são básicos para assegurar a qualidade da preservação digital a longo prazo.

Estas variáveis estão organizadas nos seguintes grupos: custo, escalabilidade, interoperabilidade, segurança de dados e facilidade de programação. Compete à organização definir a ponderação a atribuir a cada uma das variáveis, de forma a permitir essa avaliação.

Para cada uma destas variáveis deve ser atribuída uma ponderação parcial de acordo com os critérios que a organização queira privilegiar. O conjunto de variáveis por grupo dará a pontuação final desse grupo. A média ponderada dos quais dará uma classificação final ao sistema de armazenamento avaliado.

1. Custo

- Custo de instalação
- Manutenção e suporte/ano
- Prazo da garantia
- Custo da solução para 1 TB
- Custo da solução para 10 TB
- Custo da solução para 100 TB
- Custo de cada nó (1 nó suporta vários discos)
- Custo de cada unidade de controlo

2. Escalabilidade

- Número máximo de nós
- Número máximo de nós por cada unidade de controlo
- Capacidade máxima de cada nó (número de discos)
- Limite máximo em TB de cada nó

3. Interoperabilidade

- Suporta nós de marcas concorrentes
- Suporta discos de marcas concorrentes
- Disponibiliza interfaces de programação
- Suporta conectividade NFS
- Suporta conectividade SMB

4. Segurança de dados

- Suporta RAID
- Suporta *backups*
- Suporta replicação remota

5. Facilidade de programação

- *Storage Virtualization*
- API de controlo do número de réplicas
- API de controlo dos processos de auditoria

5.3 Cópias de segurança

A capacidade de fazer réplicas de igual qualidade (digital para digital) significa que é possível, e recomendável, armazenar cópias de OD num ambiente diferente das armazenadas para acesso frequente. Estas cópias terão a mesma qualidade da cópia ou original arquivado, e estes só serão necessários para inspeções, para fazer novas cópias, ou para efetuar a migração para novos formatos.

As cópias digitais de arquivo devem também ser guardadas em locais diferentes das cópias de acesso frequente, a fim de prevenir perdas de informação por desastres, roubo ou mau manuseamento.

Assim, aconselha-se a utilização de uma solução fora de linha, que limite o acesso apenas quando a informação armazenada estiver corrompida.

Cópia arquivada - uso limitado (<i>backup</i> /réplica)	Cópia de acesso frequente - uso múltiplo (armazenamento)
<ol style="list-style-type: none"> 1. Armazenar num ambiente controlado e separado, se possível, da cópia de acesso frequente; 2. Armazenar em condições ambientais recomendadas para armazenamento de arquivo. 	<ol style="list-style-type: none"> 1. Armazenar para fácil acesso; 2. Verificar, visualmente, se o disco foi alvo de algum dano após a utilização; 3. Armazenar em condições ambientais similares às condições de utilização.

O armazenamento de *backups* deve seguir um dos modelos identificados na coluna Modelo de repositórios de backups da FRD Fase 4A - Caract. Tec. Sistemas.

Unidades de suporte para *backups* mais aconselhadas:

Tapes (magnéticas): são um suporte de acesso sequencial, pelo que, mesmo que os tempos de acesso sejam baixos, a velocidade de escrita ou leitura contínua de dados é rápida. São, desde há muito tempo, o suporte mais utilizado para armazenamento, cópias de segurança, arquivo e transmissão. Existe uma variedade de formatos, muitos dos quais proprietários ou específicos de alguns nichos de mercado, como é o caso das *mainframes* ou de uma marca de computadores específica.

Discos rígidos: podem ser ligados localmente através de interfaces *SCSI*, *USB* ou *Firewire*, ou através de tecnologias de maior distância, tais como *Ethernet* (Rede), *SCSI* ou Canais de Fibra. As suas principais vantagens são: tempos de acesso baixos, disponibilidade, capacidade e facilidade de uso. A relação capacidade/preço dos discos rígidos tem vindo a melhorar ao longo dos anos.

Cópia de Segurança Remota: são feitas através Internet para um local remoto e permitem proteger os dados contra alguns dos maiores perigos, como o fogo, inundações, sismos ou detonação nuclear. À medida que as ligações de Internet de banda larga se tornam cada vez mais disseminadas, os serviços de cópia de segurança remota são cada vez mais utilizados. Um dos aspetos menos positivos desta solução, prende-se com o facto da velocidade da ligação de Internet ser normalmente mais baixa do que a velocidade de transmissão dos sistemas de armazenamento de dados locais. Este facto pode ser problemático quando se lida com grandes volumes de dados. Adicionalmente, é necessário ter em conta os perigos associados à subcontratação (*outsourcing*) de terceiros para armazenarem dados considerados sensíveis ou pessoais.

As unidades de suporte para *backups* menos aconselhadas são:

Discos Óticos: podem ser utilizados como suporte de cópias de segurança. Uma das vantagens do CD e do DVD é poderem ser restaurados em qualquer máquina com um leitor. Muitos formatos de disco ótico são *WORM*, o que faz com que sejam tradicionalmente utilizados para arquivo, visto os dados não poderem ser alterados. Outros formatos regraváveis podem também ser utilizados, como os *CDRW* ou *DVD-RAM*.

Os novos discos *HD-DVD* e *BlueRay* aumentaram de forma significativa a quantidade de dados que podem ser armazenados num único disco ótico. No entanto é preciso ter em consideração a inexistência de estudos sobre a real perdurabilidade de suportes óticos.

Disquetes: muito utilizadas durante a década de 80 e inícios de 90 e associadas às cópias de segurança. Hoje em dia, a capacidade das disquetes tornaram esta solução obsoleta.

Dispositivos de memória: também conhecidos por memórias *flash*, *Pen-drives*, cartões de memória, cartões digitais, etc., estes suportes são relativamente caros em comparação com a sua baixa capacidade, mas são bastante portáteis e fáceis de utilizar.

Nos casos em que os organismos produzam e/ou girem pequenos volumes de dados (< 10 GB) não compensa investir em grandes sistemas de armazenamento e de segurança de dados, mais vocacionados para o armazenamento de grandes volumes de dados. Nesse sentido, e tendo em conta as características específicas e limitações derivadas das ameaças referidas no capítulo 3, esses organismos poderão utilizar, após análise cuidadosa de cada uma das hipóteses de suportes de armazenamento e decisão bem justificada relativa à sua utilização como sistema de armazenamento e/ou réplica de segurança, Discos Rígidos Portáteis e/ou Discos Óticos *CD-R* e *DVD-R* e *DVD+R*. Há ainda no mercado outros formatos alternativos para os quais não se dispõe à data de produção da presente versão deste documento, de testes comprovativos das suas reais capacidades a nível de perdurabilidade.

5.3.1 Gestão do repositório de dados

Independentemente do modelo de repositório ou suporte utilizado para as cópias de segurança, é necessário estabelecer um equilíbrio entre as questões de acessibilidade, segurança e custos.

Em linha (*Online*)

Normalmente o tipo de armazenamento de dados mais acessível, que pode iniciar o restauro em milissegundos. Um exemplo disso é um disco rígido interno ou um conjunto de discos rígidos (ligados a uma SAN). Este tipo de armazenamento é muito conveniente e rápido, mas relativamente caro. Por outro lado, sofre de alguma vulnerabilidade, pois os dados podem ser apagados ou alterados, seja por acidente ou por vírus que elimine dados.

Perto de Linha (*Near-line*)

Normalmente trata-se de um tipo de armazenamento de dados menos acessível, mas mais barato, do que o armazenamento em linha, sendo útil para o armazenamento de dados de réplicas (cópias de segurança). Um exemplo disso é um conjunto de tapes com tempos de restauro que vão desde segundos até alguns minutos. Normalmente está associado a um aparelho mecânico que move as unidades de suporte do local de armazenamento para o dispositivo de leitura e/ou gravação de dados.

Fora de Linha (*Off-line*)

Similar ao perto de linha, embora necessite de intervenção humana para disponibilizar as unidades de suporte. Isto pode incluir simplesmente o armazenamento das tapes de cópias de segurança nas estantes. O tempo de acesso deste suporte pode ser muito demorado.

Cofre-Forte (*Off-site vault*)

Com vista à proteção contra desastres naturais, algumas instituições preferem enviar os suportes de armazenamento das cópias de segurança para um cofre-forte fora de linha. O cofre pode ser simplesmente o gabinete do Administrador do sistema ou, algo tão sofisticado como, um *bunker* de alta segurança com sistemas anti-desastre e controlo climatérico, que permitem o armazenamento de suportes de cópias de segurança.

Unidade de Cópias de Segurança/Centros de Recuperação de Dados (*Backup site, Disaster Recovery Center*)

No caso de um desastre, os dados nos suportes de cópia de segurança não serão suficientes para a recuperação da informação. Também são necessários sistemas informáticos e redes bem configuradas para os dados poderem ser restaurados. Algumas organizações têm os seus próprios centros de recuperação de dados equipados para este tipo de cenário. De notar que, sendo estas unidades de recuperação de dados um investimento tão dispendioso, raramente se utilizam as cópias de segurança como método de transferência de dados para a unidade de recuperação de dados. O procedimento mais usual é a replicação de disco (*disk mirroring*) remota, que permite manter os dados no centro de recuperação de dados o mais atualizado possível.

6 Metainformação

Este capítulo aborda a problemática da necessidade de metainformação (MI) para uma correta e viável preservação digital. Para tal refere-se a utilização de várias camadas de MI, como forma de a organizar, indicando-se, para cada camada, exemplos de esquemas de MI passíveis de serem utilizados. É proposta uma lista de elementos de MI recomendados para a camada de MI de preservação. Finalmente, alerta-se para a utilização de uma linguagem cuja sintaxe permita ser legível e pesquisável pelo ser humano e pelas máquinas.

A Metainformação - ou metadados - é transversal a todos os elementos (identificados nos Cap. 4 e 5) deste documento.

Na aceção mais básica, a MI é informação sobre informação, ou mais especificamente dados estruturados sobre informação capturada no sistema de arquivo.

A metainformação descreve atributos do documento de arquivo Eletrónicos dando-lhe significado, contexto e organização, permitindo a produção, gestão e utilização de documentos de arquivo ao longo do tempo, assim como nos, e através dos, domínios em que são produzidos.

Note-se que existem muitos esquemas de metainformação - normalizados ou não - destinados a descrever aspetos temáticos ou funcionalmente especializados. Por exemplo existe metainformação específica para dados estatísticos ou para sistemas espaciais como os SIG.

A metainformação pode ainda existir implicitamente embebida nos objetos digitais sendo automaticamente produzida pelas soluções aplicacionais que lhes deram origem

No âmbito da documentação de arquivo eletrónica, a MI consiste num conjunto de elementos (atributos) associados a cada OD no momento da sua integração no arquivo e ao longo da sua gestão.

A metainformação pode ser associada não apenas a documentos individuais mas também a pastas, utilizadores, plano de classificação e qualquer outra entidade de um Sistema de Gestão Documental/Arquivo.

A importância da MI neste âmbito prende-se com a dependência dos documentos eletrónicos relativamente ao sistema intermediário, sendo então necessária para que os OD fiquem contextualizados em termos de produção, gestão e preservação. A metainformação associada aos OD permite:

- A referenciação única e persistente de cada OD;
- A localização e recuperação expedita de cada OD;
- A criação do histórico de cada OD;
- O registo das transformações realizadas sobre cada OD;

Embora de uma forma geral os documentos eletrónicos contenham alguma metainformação fornecida pelas aplicações que os produzem, esta é insuficiente e/ou não se encontra estruturada (organizada). Para ultrapassar esse facto dispomos de esquemas de metainformação que designam os elementos que têm de ser associados aos documentos eletrónicos. Importa clarificar que não existem esquemas de MI obrigatórios, pese embora a sua utilização se inclua no conjunto de boas práticas para a preservação digital.

Visto que a MI pode ser repartida por áreas, é aconselhável a utilização de diferentes esquemas no âmbito da preservação digital, que devem ser aplicados por diversas camadas:

■ **Metainformação Descritiva ou de Identificação:**

O objetivo é a pesquisa, recuperação e identificação. Pode incluir elementos como o título, o assunto, autor e palavras-chave. O esquema EAD (*Encoded Archival Description*) é um bom exemplo, por ser um esquema muito completo e que respeita as normas de descrição arquivística;

O sítio Web da Library of Congress inclui a página oficial do EAD: <URL:<http://www.loc.gov/ead>>.

O RODA utiliza um modelo de dados baseado no EAD, designado por EADPART, abordado em <URL:<http://roda.iantt.pt/pt/system/files/roda-relatorioFinal.pdf>>

■ **MI Administrativa:**

Fornecer informação para apoio à gestão do documento de arquivo Eletrónicos, tal como quando e como foi criado, tipo de ficheiro e outra informação técnica, e quem tem privilégios de acesso;

■ **MI Estrutural:**

Relaciona hierarquicamente os diferentes OD que fazem parte de um mesmo documento. O esquema METS (*Metadata Encoding & Transmission Standard*) permite agrupar metainformação descritiva, administrativa e estrutural sobre objetos guardados num repositório digital.

O sítio Web da Library of Congress inclui a página oficial do METS:

<URL: <http://www.loc.gov/standards/mets>>.

O esquema METS foi escolhido no âmbito do RODA, por permitir a codificação da MI relativa ao OD, nomeadamente a sua estrutura hierárquica, o seu nome e localização.

■ **MI Técnica:**

Descreve as características técnicas dos ficheiros e dos seus formatos, pelo que é a mais específica, por depender do formato do OD. Um exemplo deste tipo de MI é o esquema baseado na norma ANSI/NISO Z39.87 - *Metadata for Images in XML schema*, define um conjunto normalizado de elementos de metainformação para imagens digitais;

O sítio Web da Library of Congress inclui a página oficial da ANSI/NISO Z39.87: <URL:<http://www.loc.gov/standards/mix>>

O RODA utiliza o esquema baseado na norma ANSI/NISO Z39.87 no caso das imagens digitais.

■ **MI de preservação:**

Contém informação necessária para arquivar e preservar o OD. O esquema PREMIS (*PREservation Metadata Implementation Strategies*), produto da parceria entre a OCLC (*Online Computer Library Center*) e a RLG (*Research Libraries Group*) é o mais conhecido,

O sítio Web da Library of Congress inclui a página oficial do PREMIS: <URL:<http://www.loc.gov/standards/premis>>

O RODA definiu o PREMIS como o esquema MI de preservação a ser utilizado no âmbito do projeto.

Existem, adicionalmente, algumas subcamadas de tipos de metainformação, que são aqui identificadas em separado:

- MI de gestão de controlo e direitos: que tem a ver com os direitos de propriedade intelectual;
- MI de contexto de produção;
- MI de controlo e gestão.

O organismo deve fazer uma descrição de todos os esquemas, modelos ou elementos de MI utilizados ou que lhes interesse utilizar, considerados suficientes. Para tal deve produzir uma tabela que contenha essa informação.

É preciso ter em conta que, no contexto da preservação digital, existem elementos de MI que devem ser considerados obrigatórios. Para tal, é avançada, no quadro abaixo, uma proposta que especifica os elementos recomendados, retirados do esquema *PREMIS*⁷. A primeira coluna indica o sujeito do conjunto dos atributos; a coluna ID apresenta o número do atributo de acordo com a ordenação existente no dicionário de dados do *PREMIS*; a coluna atributo identifica as unidades semânticas; a coluna Descrição estabelece o âmbito dos atributos.

⁷ O Esquema PREMIS está disponível em <http://www.oclc.org/research/projects/pmwg/>

	Id.	Atributo (<i>semantic unit</i>)	Descrição
Objeto	1	• {objectIdentifier}	Uma designação para identificar de forma unívoca um objeto num arquivo digital.
	1.1	• <i>objectIdentifierType</i>	
	1.2	• <i>objectIdentifierValue</i>	O valor do identificador do objeto.
	2	• <i>preservationLevel</i>	Um valor a indicar o conjunto de funcionalidades de preservação a aplicar ao objeto. Pode ser adequado consoante aquilo que se pretende que o OD assegure.
	3	• <i>objectCategory</i>	A categoria de objeto a que a MI é aplicável. A profundidade de MI aplicável, de acordo com o elemento 2.
	4	• {objectCharacteristics}	As propriedades técnicas de um ficheiro ou <i>bitstream</i>
	4.1	• <i>compositionLevel</i>	Indicação sobre se o OD está sujeito a uma ou mais operações de descodificação ou decomposição. No caso do OD estar comprimido ou agregado a outros OD e ser necessário isolá-lo.
	4.4	• {format}	Organização de um ficheiro ou <i>bitstream</i> em que se entende por formato a organização da informação digital de acordo com condições pré-definidas. Registrar o formato para cada ficheiro que compõem o OD.
	4.4.1.1	• <i>formatName</i>	Designação do formato do ficheiro ou <i>bitstream</i> .
	7	• {storage}	Informação sobre como e onde o OD é armazenado no Arquivo Digital.
	7.1.1	• <i>contentLocationType</i>	Esquema de utilização utilizado (os meios para aceder à localização).
	7.1.2	• <i>contentLocationValue</i>	A referência à localização do conteúdo.
	7.2	• <i>storageMedium</i>	O suporte físico em que o objeto está armazenado.
Evento	14	• {eventIdentifier}	Uma designação para identificar de forma única um evento decorrido dentro do Arquivo Digital.
	14.1	• <i>eventIdentifierType</i>	Designação para o domínio em que o Identificador do evento é único (ex. sistema de armazenamento, integração, monitorização, etc. ou um identificador único sequencial para todo o Arquivo Digital).
	14.2	• <i>eventIdentifierValue</i>	Valor do identificador.
	15	• <i>eventType</i>	Categorização da natureza do evento (de acordo com as funcionalidades asseguradas pelo Arquivo Digital e que devem constar de uma lista de autoridade).
	16	• <i>eventDateTime</i>	A data ou intervalo de datas em que o evento decorreu.
Actor	21	• {agentIdentifier}	A designação usada para identificar de forma única um ator dentro de um sistema de arquivo digital.
	21.1	• <i>agentIdentifierType</i>	Designação do domínio em que o id do ator é único.
	21.2	• <i>agentIdentifierValue</i>	Valor do id do ator.
Direitos	28	• {permissionGranted}	As ações permitidas pelo depositante ao AD.
	28.1	• <i>act</i>	Ação que o AD pode empreender.
	28.3.1	• <i>startDate</i>	Data de início do acordo.
	28.3.2	• <i>endDate</i>	Data em que termina o acordo.

A metainformação estruturada deve basear-se numa sintaxe que facilite a pesquisa e a identificação por parte tanto do utilizador como da máquina. A sintaxe mais utilizada na atualidade é o XML (eXtensible Markup Language), desenvolvida pelo W3C (World Wide Web Consortium), uma forma estendida de HTML

O Arquivo Nacional da Holanda produziu um *Digital Preservation Testbed White Paper* sobre o XML
<www.digitaleduurzaamheid.nl/bibliotheek/docs/white-paper_xml-en.pdf>

que permite a definição localizada de etiquetas e a troca facilitada de informação. Atualmente, existem cada vez mais ferramentas para XML, e esta linguagem tem assumido um papel cada vez mais crucial na troca de uma variedade de dados na Web.

Apresentamos um exemplo da utilização do XML no **Anexo C**.

7 Estratégias de Preservação mais comuns

Este capítulo identifica as estratégias de preservação mais utilizadas, a saber, preservação de tecnologia, emulação, monitorização de suportes e formatos, encapsulamento e a migração e transferência de suportes. Neste âmbito, são referidas as características, bem como os aspetos positivos e negativos de cada uma destas estratégias.

A seleção da(s) estratégia(s) de preservação apropriada(s) deve ser, como já foi referido, o resultado de um esforço de colaboração entre as unidades orgânicas referentes ao arquivo e à informática, com a participação de todas as unidades orgânicas afetadas pelo processo ou que produzam documentos de arquivo eletrónicos.

Apresenta-se, seguidamente, o conjunto das estratégias de preservação mais comumente utilizadas, elencando características, vantagens e desvantagens de cada uma delas. As estratégias apresentadas podem ser utilizadas a curto, médio e longo prazo, cabendo à entidade produtora definir a sua calendarização e utilização no seu PPD

7.1 Preservação de tecnologia

A preservação de tecnologia implica a conservação e manutenção de todo o *hardware* e *software* necessários à correta apresentação dos OD. Neste âmbito, o foco da preservação concentra-se, não no objeto conceptual, mas sim na preservação do objeto digital na sua forma original.

A preservação de tecnologia apresenta algumas desvantagens, nomeadamente:

- O facto de qualquer plataforma tecnológica acabar por se tornar obsoleta;
- A existência de dificuldades na gestão do espaço físico;
- A manutenção e custos de operação;
- O facto do acesso à informação estar limitado a alguns locais físicos do globo e com condicionalismos acrescidos ao nível da reutilização da informação.

7.2 Emulação

A emulação corresponde à utilização de um *software* - o emulador - capaz de reproduzir o comportamento de uma plataforma de *hardware* e/ou *software*, numa outra plataforma que, à partida, seria incompatível.

As principais vantagens da emulação incluem o facto de preservar, de forma fiel, as características e as funcionalidades do objeto digital original e, embora se centre na preservação do objeto lógico no seu formato original, não

Tipos de emulação possíveis:

Emulação de aplicações informáticas (ex: *Lotus 123*);
Emulação de sistemas operativos (ex: emulador de MS Windows para Linux);
Emulação de plataformas de hardware (ex: emulador de ZX Spectrum para PC).

Jeff Rothenberg, defensor da emulação, propõe um modelo teórico passível de emular plataformas atuais em computadores futuros, através da conservação do objeto digital juntamente com o software necessário à sua execução/apresentação e na criação de uma especificação escrita numa linguagem independente da plataforma de hardware que suporta a execução desse software.

www.clir.org/pubs/reports/rothenberg/pub77.pdf

sofre de alguns dos problemas da estratégia de preservação de tecnologia (Ex.: envelhecimento do hardware).

No entanto, o emulador pode tornar-se obsoleto e a sua utilização pressupõe que os utilizadores do futuro sejam capazes de operar adequadamente aplicações e sistemas operativos há muito desaparecidos (Ex.: num futuro próximo será difícil conceber que os utilizadores estejam aptos a enfrentar as particularidades do sistema operativo MS-DOS).

T. Hendley defende que a emulação apenas deveria ser utilizada em situações em que haja interesse em preservar o ambiente tecnológico original ou quando não é possível converter os OD para formatos atuais.

David Bearman considera que a emulação pretende conservar o objeto errado, i.e., as funcionalidades dos sistemas de informação em vez dos DAE, pois são estes últimos que têm valor evidencial, através do seu conteúdo, estrutura e contexto.

O Arquivo Nacional dos Países Baixos produziu um *Digital Preservation Testbed White Paper* sobre a Emulação
<www.digitaleduurzaamheid.nl/bibliotheek/docs/white_paper_emulatie_EN.pdf>

7.3 Monitorização de suportes e formatos

A monitorização prevê processos de verificação automática, manual e semiautomática dos OD, sendo as duas últimas opções mais realistas por razões de custos.

As preocupações da monitorização prendem-se com as seguintes questões:

- Tempo de vida estimado dos suportes;
- Tempo médio de prevalência de uma versão de aplicação informática (3 anos);
- Retrocompatibilidade assegurada pelos fabricantes (em média 3 versões anteriores).

7.4 Encapsulamento

O encapsulamento consiste em preservar, juntamente com o objeto digital, toda a informação necessária e suficiente para permitir o futuro desenvolvimento de conversores, visualizadores ou emuladores (por exemplo, a descrição formal e detalhada do formato do objeto preservado).

O encapsulamento está orientado a objetos que apenas serão acedidos num futuro longínquo. Permite adiar a responsabilidade de preservação, e o desenvolvimento futuro de visualizadores, migradores ou emuladores.

No entanto, é preciso estar ciente que objetos complexos possuem especificações complexas e que uma especificação incompleta poderá ter um efeito desastroso para a preservação do OD.

7.5 Transposição de Formatos e Suportes (Migração e transferência de suporte)

A transposição de formatos e suportes refere-se à transferência de documentos contidos num determinado suporte ou formato para outro suporte ou formato mais atualizado. É o processo responsável pela reorganização dos elementos de informação que constituem um OD. Centra-se, sobretudo, na preservação do seu conteúdo intelectual, ou seja, na preservação do objeto conceptual e na verificação frequente da integridade dos suportes físicos.

O principal objetivo é evitar a obsolescência tecnológica, mantendo os OD compatíveis com tecnologias atuais, de forma a permitir a sua interpretação sem necessidade de recorrer a artefactos menos convencionais.

A transposição prevê processos de:

- Refrescamento a nível de suportes (exemplo: de Disquete para CD-R ou de CD-R para DVD-R), nos casos em que existam e sejam utilizados estes suportes (cf. Cap. 5).
- Processos de migração entre formatos (exemplo: de Word 97 para Word 2000);
- Transposição conjunta de formatos e suportes.

O Arquivo Nacional dos Países Baixos produziu um *Digital Preservation Testbed White Paper* sobre a Migração <www.digitaleduurzaamheid.nl/bibliotheek/docs/Migration.pdf>

Esta estratégia apresenta, no entanto, algumas desvantagens como sejam a:

- Probabilidade de algumas propriedades dos OD não serem corretamente transferidas para o formato de destino adotado;
- Existência de incompatibilidades entre os formatos de origem e destino;
- Inadequação dos conversores;
- Obsolescência de formatos.

Adicionalmente, durante a transposição de documentos para novos formatos há informação que se perde ao nível da estrutura, da metainformação e, por vezes, do conteúdo.

Face a estas situações de perda de informação, deve recorrer-se a estratégias nas quais se incluem documentar exaustivamente o processo de migração (metainformação) para que fique claro o que se perdeu, quando e como se perdeu. Desta forma tenta-se garantir que o que se perde não compromete a autenticidade e fidedignidade do documento. Este aspecto de confirmação da perda de informação em objetos migrados foi já referido no capítulo 2.

Esta questão é particularmente evidente em documentos compostos e complexos como, por exemplo:

- Bases de dados (muitas tabelas com ligações entre si e circulação constante de informação);
- Documentos multimédia (páginas *web* em que as atualizações e modificações são frequentes).



Apesar das desvantagens apresentadas, esta é de longe a estratégia de preservação mais aplicada até à data e a única que tem vindo a dar provas da sua eficácia.

Existem algumas variantes no caso específico da transferência de formatos, a saber:

■ Atualização de versões:

Caracteriza-se pela atualização dos formatos utilizando *software* retro compatível (por exemplo, a utilização do Word 6 para atualizar um documento criado com o Word 5);

■ **Conversão para formatos concorrentes:**

Permite ultrapassar o risco de descontinuidade de formatos, convertendo OD para formatos análogos, independentemente da aplicação utilizada na sua criação (por exemplo, converter um documento Word para PDF);

■ **Normalização:**

Corresponde à migração para um número reduzido de formatos compatíveis, o que poderá evitar futuras complicações a nível de direitos de autor ou pagamento de *royalties*. Promove, também, a interoperabilidade entre sistemas distintos.

O facto de serem utilizados formatos abertos e independentes da plataforma permite que diferentes configurações de *hardware* e *software* sejam capazes de os interpretar. Existem vários tipos de normas:

● **Normas *de facto*:**

Quando o produto está firmemente estabelecido no mercado, a compatibilidade das outras aplicações é medida referenciando esse produto e as alterações de compatibilidade dependem do proprietário (ex: Word);

● **Especificações públicas:**

Quando algumas empresas líderes de mercado criam um consórcio para definir uma norma de ligação que possibilita o desenvolvimento de produtos compatíveis;

● **Normas *de jure***

Quando organismos oficiais promovem o consenso acerca de uma especificação, que depois se torna uma norma oficial (ex: ISO, NP, etc.)

O tempo de intervalo entre transferências de formato está dependente da evolução da tecnologia, bem como do facto de usarmos ou não formatos normalizados, podendo variar entre os 7 e os 10 anos.

8 Conclusão

A preservação digital é uma função que não pode ser evitada. Cada vez mais as organizações dependem, a médio e longo prazo, de informação produzida e mantida eletronicamente. Tal dependência deve-se a motivos que se prendem com a operação da organização, salvaguarda dos seus interesses, incrementação da transparência administrativa e credibilidade da instituição relativamente aos seus *stakeholders*, entre os quais se contam o Governo e o cidadão.

A organização pode ter a opção de não preservar a sua informação digital. Julgamos, no entanto, ser esse risco inaceitável atendendo às circunstâncias referidas. Embora tal implique custos que podem ser significativos, compete à organização ponderar seriamente o que tem a perder caso decida nada fazer.

Este documento é um primeiro passo para guiar as organizações na tarefa de preservação digital. A DGARQ, enquanto órgão responsável pela coordenação e execução da política arquivística nacional, posiciona-se como parte interessada nesse processo, disponibilizando o seu conhecimento na área para auxiliar as organizações a construir e implementar processos de preservação, sem se pretender substituir de forma alguma às responsabilidades que toda e qualquer organização deve manter relativamente ao seu património arquivístico, independentemente do suporte em que este se encontra.

Anexos

Instruções de utilização do Anexo A - Lista de requisitos para a prática de preservação digital

O Anexo A é:

Uma *check-list* dos requisitos que têm que ser cumprido pelas Organização que pretendam produzir um Plano de Preservação Digital

Tem como objetivos:

- Indicar se a Organização está preparada para a preservação de DAE;
- Indicar os requisitos a cumprir para a Organização se considerar preparada para a prática da preservação digital

Deverá ser utilizado quando:

- no início do projeto de planificação, para saber quais as condições existentes na Organização no âmbito da preservação digital;
- durante o projeto, como indicador do estado da Organização relativamente à planificação do Plano de Preservação Digital;
- No fim da planificação do projeto, para verificar se todos os requisitos foram tidos em conta, ou seja, se todos os requisitos considerados como obrigatórios ou, obrigatórios se aplicáveis, foram cumpridos.

NOTA: *A ordem dos pontos apresentada no anexo não terá que ser cumprida sequencialmente*

Ponto 1

Os requisitos do ponto 1 referentes ao Plano de Classificação e à Tabela de Seleção dizem respeito à existência destes documentos atualizados por parte da Organização.

Ponto 2

[2.1] Um **sistema de arquivo** é um sistema de informação que integra, gere e fornece acesso a documentos de arquivo ao longo do tempo. Nessa acepção, a **gestão documental** tem como objetivo o controlo eficiente e sistemático da produção, recepção, manutenção, utilização e destino dos documentos de arquivo, incluindo os processos para constituir e manter prova e informação sobre atividades e transações. Isto inclui qualquer prática relacionada com a documentação, e que poderá ir desde a simples numeração, ordenação e/ou identificação dos documentos, até práticas mais avançadas como a utilização de sistemas eletrónicos de gestão de documentos.

A **norma NP 4438** dá orientações relativas à gestão de documentos de arquivo (exclui documentos históricos) nas entidades produtoras, públicas ou privadas, para utilizadores internos ou externos. A DGARQ fornece consultoria nesta área



Um sistema de gestão documental não implica a existência de uma aplicação informática de gestão documental. No entanto, **para fins de preservação** da autenticidade, fidedignidade, integridade e usabilidade dos documentos de arquivo eletrónicos, **supõe-se a utilização de SEGAs.**

[2.2] Uma **série documental** (independentemente do suporte) é um conjunto de documentos que obedece a idêntica competência funcional e tipológica. Neste contexto pode tratar-se de um:

- Sistema de informação que suporte uma determinada atividade. [**Classe de Informação**];
- Conjunto de sistemas de informação que suporte a mesma função ou atividade [**atividades suportadas**], ou ainda
- Conjunto de documentos eletrónicos não estruturados (por exemplo documentos de texto e mensagens de correio Eletrónicos que façam parte de um processo administrativo).

§ Um sistema de informação pode contemplar várias séries (ex: uma base de dados pode incluir informação referente a várias séries documentais).

Exemplo de SI: Sistema de Gestão de Recurso Humanos

Exemplos de Atividades Suportadas pelo SI: Avaliação e Desempenho; Processos, Registos e Meios de Identificação; Prestação, Suspensão e Cessação de Trabalho; Acidentes em Serviço; Formação; Remunerações, Abonos e Descontos;

Exemplos de Séries contempladas:

Avaliação e Desempenho: Processos de avaliação do desempenho individual; Processos de reclamação e recurso da avaliação

Registos e Meios de Identificação: Processos individuais de funcionários, Cartões de identificação

Prestação, Suspensão e Cessação de Trabalho: Processos de marcação de férias, Registos de assiduidade, Escalas de pessoal, Processos de verificação de doença e junta médica, Processos de pedidos de aposentação, Outras situações de cessação jurídica de emprego

Formação: Plano e relatório anual de formação interna, Processos de ações de formação, Estágios, Bolsa de formadores, Divulgação de ações de formação

Exemplo de SI: Sistema Informação Contabilística

Exemplos de Funções Suportadas: Património, Instalações e Recursos Materiais; Gestão Orçamental

Exemplos de Atividades Suportadas pelo SI:

Património, Instalações e Recursos Materiais: Aprovisionamento; Gestão de bens móveis; Gestão de Imóveis; Gestão de Stocks

Gestão Orçamental: Orçamento de Estado; Alterações Orçamentais; Controlo da execução orçamental

Exemplos de Séries contempladas:

Património, Instalações e Recursos Materiais:

Aprovisionamento: Aquisição de bens e serviços; Relatórios de controlo da execução dos contratos; Ficheiros de fornecedores

Gestão de bens móveis: Inventário; Abate de bens móveis; Cedência de bens móveis; Manutenção de equipamentos

Gestão de Imóveis: Inventário; Aquisição/ alienação; Arrendamento; Processos de cedência e aluguer de espaços/ instalações

Gestão de Stocks: Mapas de existências; Requisições de material/ equipamento; Guias de entrada e saída de material/ equipamento

Gestão Orçamental:

Orçamento de Estado: Orçamento de funcionamento; Orçamento PIDDAC

Alterações Orçamentais: Alterações orçamentais; Pedidos de autorização de pagamentos; Pedidos de libertação de crédito; Controlo do movimento bancário

Controlo da execução orçamental: Relatórios de execução; Despesas cofinanciadas; Despesas relativas a programas institucionais

Ponto 3

[3.1] As séries documentais/classes de atividades suportadas/classes de informação que devem ser preservadas são:

- as que contêm informação digital, e com
- prazo de conservação superior a 7/10 anos.

[3.2] As características de cada tipo de ficheiro que permitem descrever a experiência visual e contextual dos utilizadores ao lidar com os OD) são:

- Forma (aparência) - o modo como a informação é reproduzida (por ex., como matéria impressa, como apresentação gráfica bidimensional, sob forma exclusivamente textual, como discurso sintetizado, em Braille);
- Conteúdo - aquilo que a informação transmite ao utilizador por meio de linguagem natural, imagens, sons, filmes, animações;
- Estrutura - o modo como a informação está organizada em termos lógicos (por exemplo, por capítulos, com ou sem uma introdução e um sumário).;
- Comportamento - conjunto de reações programadas do ficheiro face às interações do utilizador ou do sistema.

Exemplo: Base de dados

Forma: dependente do formulário ou relatório;

Conteúdo: dados colocados via formulário;

Estrutura: tabular;

Comportamento: atualização de data e hora (caso esteja assim definido nas características).

Exemplo: Ficheiro PDF

Forma: Texto Estruturado;

Conteúdo: Ofício;

Estrutura: paginado;

Comportamento: Estático

Exemplo: Folha de Cálculo Microsoft Excel

Forma: Tabelas;

Conteúdo: Cálculos;

Estrutura: Tabular;

Comportamento: Atualiza datas, atualiza juros ...

Exemplo: Vídeo em formato AVI

Forma: Imagem em movimento e som

Conteúdo: Vídeo e Áudio;

Estrutura: Por capítulos;

Comportamento: Dependente dos comandos do utilizador

Deve-se, ainda, ter em conta a discriminação das características específicas que um determinado formato possa ter. Por exemplo nos ficheiros de imagem poderá incluir-se

informação acerca da profundidade de cor, resolução de imagem, tipos de interatividade, possibilidade de utilização de macros e metainformação embebida;

[3.3] As características da informação de arquivo eletrónica incluem:

- o contexto: O contexto de produção do documento/série/classe de informação;
- o suporte de armazenamento: Material adequado ao registo de informação (CD, DVD, tape, discos rígidos);
- Relações existentes com outros documentos de arquivo;
- Metainformação que lhes está associada ou é atribuída pela organização);

Contexto: Justificação de faltas;
Suporte de armazenamento: servidor;
Relações existentes com outros documentos de arquivo: inclui informação contida na base de dados do telepono; faz parte do processo individual do funcionário;
Metainformação: Assunto, Data, Colaborador, Data da Falta, Justificação, Despacho.

Contexto: Registo Eletrónico de Documentos analógicos;
Suporte de armazenamento: servidor;
Relações existentes com outros documentos de arquivo: inclui informação sobre todos os documentos de arquivo em suporte analógico
Metainformação: Assunto, Data, Produtor, Destinatário, Classificação, Nº de Ref., Localização do documento, Informação/Despacho.

Contexto: Pedido de Consultoria (email);
Suporte de armazenamento: disco rígido;
Relações existentes com outros documentos de arquivo: Documento que inicia o processo de consultoria.
Metainformação: Assunto, Data, Produtor, Destinatário, Classificação, Nº de Ref., Localização do documento, Informação/Despacho.

[3.4] Os fatores inerentes à informação de arquivo eletrónica que irão influenciar opções a incluir no Plano de Preservação Digital são, por exemplo:

- A usabilidade - capacidade da informação de arquivo ser utilizável ou seja, que pode ser localizada, recuperada, apresentada e interpretada;
- A escalabilidade - característica de um sistema ou equipamento que pode crescer em escala, isto é, que possibilita incrementos de capacidade ou funcionalidades acompanhando as necessidades dos utilizadores (por exemplo o aumento da capacidade de armazenamento do sistema através da substituição ou instalação de novos discos rígidos; versões mais atuais do formato PDF que permitem a inclusão de informação de dados CAD);
- A complexidade significa que a informação de arquivo (um OD) integra diversas componentes digitais (por exemplo as páginas *Web*);

Exemplo: Documento Arquivo Eletrónico em formato PDF
Usabilidade: Possibilidade de visualização e impressão. Não editável;
Escalabilidade: PDF v1.9, Escalável para versões posteriores existentes
Complexidade: Contém Texto estruturado, Imagem Fixa Digitalizada, Áudio e Vídeo e Gráficos tridimensionais.

Exemplo: Sistema de Assiduidade (Relógio de Ponto Digital)
Usabilidade: Possibilidade de visualização e impressão. Não editável;
Escalabilidade: Sistema em código aberto. Hardware permite aumento de memória (RAM), capacidade de armazenamento, e velocidade de processamento (CPU).
Complexidade: Contém dados tabulares (dados registados) e ficheiros do sistema . Permite a produção de relatórios em formato PDF.

Exemplo: Sítio Web
Usabilidade: Possibilidade de visualização e impressão. Informação editável;
Escalabilidade: Sistema em código aberto. Hardware permite aumento de memória (RAM), capacidade de armazenamento, e velocidade de processamento (CPU).
Complexidade: Contém dados tabulares (base de dados) e ficheiros do sistema em formato HTML, CSS, PHP. Contém ainda Imagem Fixa Digitalizada, Áudio e Vídeo

Ponto 4

A avaliação e a escolha destes pontos requerem conhecimentos técnicos aprofundados a nível de informática e das novas tendências ao nível da preservação digital.

[4.1] **Formatos normalizados** - Formatos cujas definições e características estão especificadas numa norma, ou são consensualmente utilizado pela comunidade cibernética. Estes devem ser utilizados para a prática da preservação digital, para todos os formatos de ficheiros que se verificou necessário preservar;

Formato de Texto Estruturado
Formatos Utilizados: MS Word, MS Excel, RTF, TXT, EML LaTeX
Formato Normalizado: PDF/A (PDF v1.4)

Formato de Imagem Fixa
Formatos Utilizados: JPEG, BMP, PNG, PS; GIF
Formato Normalizado: TIFF

Formato de Base de Dados
Formatos Utilizados: MS Access, MySQL, MS SQL Server, PostgreSQL
Formato Normalizado: DBML (XML)

[4.2] **Aplicações informáticas** - Programa concebido como ferramenta auxiliar para o desempenho de uma determinada tarefa, como processamento de texto, a contabilidade ou a gestão de inventários

- **Incremento de utilizadores simultâneos** - capacidade de aumentar o número de utilizadores ao mesmo tempo
- **Formato aberto** - Formato com especificações públicas para armazenamento, geralmente mantido por organizações de normas não proprietárias e de uso livre de restrições legais.
- **Exportação para formatos normalizados** - Vide 4.1

[4.3] **Sistema de armazenamento** - Dispositivo ou conjunto de dispositivos utilizados para armazenar primariamente os dados e a informação produzida

- **Escalamento da capacidade de armazenamento** - característica do sistema ou equipamento que pode crescer em escala, isto é, que possibilita incrementos de capacidade ou funcionalidades acompanhando as necessidades dos usuários (aumento da capacidade de armazenamento do sistema através da substituição ou instalação de novos discos rígidos);
- **Sistema de Backup** - Sistema que permita a cópia de um programa, disco ou dados, destinada a ser arquivada ou a salvaguardar ficheiros valiosos, caso os originais sejam destruídos;
- **Sistema de Rotinas de Controlo e Monitorização** - Sistema que permite o registo de procedimentos e verificação e a identificação de eventuais disfunções e empreender ações corretivas.

Ponto 5

Este ponto diz respeito à metainformação:

- [5.1] já utilizada pelo organismo;
- [5.2] a ser utilizada no futuro.

Metainformação Descritiva - Informação produzida com objetivos de identificação, pesquisa e recuperação dos documentos/séries/classes de informação;

Metainformação Técnica - Informação que permite descrever as características técnicas dos ficheiros e dos seus formatos;

Metainformação Estrutural - Informação que permite relacionar hierarquicamente os diferentes OD que fazem parte de um mesmo documento;

Metainformação de preservação - Contém informação necessária para arquivar e preservar os OD.

Exemplo: Esquemas de MI Utilizados no SEGA

MI Descritiva: EAD

MI Técnica: ANSI/NISO Z39.87 (Imagens)

Exemplo: Esquemas de MI Utilizados no Sítio Web

MI Descritiva: Dublin Core

MI Técnica: ANSI/NISO Z39.87 (Imagens)

Exemplo: Esquemas de MI Utilizados no SI Contabilidade

MI Interoperabilidade: MIP (Metainformação para a Interoperabilidade)

Ponto 6

Este ponto tem a ver com os estudos de mercado necessários para conhecer os custos relativos à fase de produção do Plano de Preservação Digital e a sua implementação.

Estas análises podem ser feitas auscultando os fornecedores, os prestadores de serviço, as empresas de assessoria e incluir tendências a 3 anos. A existência de cadernos de encargos do que se pretende pedir aos fornecedores ajudará a obter informação mais rigorosa.

O resultado tem ser uma tomada de decisão sobre as necessidades versus custos.

Software - Pretende-se a utilização de Software livre, i.e., qualquer programa de computador que pode ser usado, copiado, estudado, modificado e redistribuído sem nenhuma restrição. A liberdade de tais diretrizes é central ao conceito, o qual se opõe ao conceito de software proprietário, mas não ao software que é vendido almejando lucro (software comercial). A maneira usual de distribuição de software livre é anexar a este uma licença de software livre, e tornar o código fonte do programa disponível. O software livre também é conhecido pelo acrónimo FLOSS (do inglês *Free/Libre Open Source Software*).

Hardware - é a parte física do sistema informático, ou seja, é o conjunto de componentes eletrónicos, circuitos integrados e placas, que se comunicam através de um conjunto de linhas de comunicação. Pretende-se preferencialmente hardware que:

- Seja fácil de instalar e administrar;
- tenha um desempenho adequado às necessidades específicas de cada organização, de forma a evitar “engarrafamentos” (bloqueios e atrasos de reposta);
- Estar em conformidade com normas *standard, de jure*;
- Seja independente da plataforma aplicacional utilizada (sistema operativo e vendedor de *hardware*);
- Seja interoperável, isto é, permita o funcionamento e comunicação com diferentes sistemas;
- Seja escalável;
- E que o custo inicial e de manutenção deverá ser mínimo, tal como o esforço necessário para a sua administração;
- Tenha garantias a nível de segurança de dados.

Infraestruturas - conjunto de elementos estruturais que enquadram e suportam todo o sistema informático, incluindo o sistema intermediário responsável pela preservação digital. Isto inclui aspetos como prevenção de falhas a nível sistema de controlo climático, eletricidade e de canalização. Devem ser tidas em conta as normas ISO/IEC 17799:2005, ISO 18934:2006, ISO/IEC 27001:2005, ISO/IEC 27002:2005, ISO/IEC 27006:2007, **CobIT** e **SEGNAC**.

Ponto 7

Este ponto tem a ver com os estudos de mercado necessários para conhecer os custos relativos ao funcionamento contínuo, manutenção e atualização, após a implementação do Plano de Preservação Digital.

Estas análises podem ser feitas auscultando os fornecedores, os prestadores de serviço, as empresas de assessoria, e podem que incluir tendências a 3 anos.

A existência de cadernos de encargos do que se pretende pedir aos fornecedores ajudará a obter informação mais rigorosa.

O resultado tem ser uma tomada de decisão sobre as necessidades versus custos.

Ponto 8

Este ponto diz respeito à alocação e afetação formal dos recursos humanos, materiais e financeiros considerados necessários. Isto implica a orçamentação do projeto, a definição de objetivos para os colaboradores e prestadores de serviços.

Ponto 9

Este ponto tem a ver com a existência um cronograma de execução das medidas necessárias para a implementação do Plano de Preservação Digital. Este cronograma deve incluir balizas cronológicas para as diversas fases do projeto.

NOTA: As respostas aos pontos 10 a 15 pressupõem já a existência de um Plano de Preservação Digital.

Ponto 10

Este ponto refere-se à definição de medidas para controlo e avaliação do funcionamento dos sistemas utilizados para a preservação digital:

- [10.1] os procedimentos de preservação;
- [10.2] os formatos para preservação;
- [10.3] soluções de software alternativas para um determinado formato;
- [10.4] soluções de armazenamento;
- [10.5] esquemas de metainformação.

Ponto 11

Este ponto refere-se à definição de medidas de monitorização interna da Informação de Arquivo Eletrónica a preservar, e que permitam verificar a manutenção:

- [11.1] da sua integridade
- [11.2] da sua funcionalidade
- [11.3] da sua estrutura
- [11.4] do seu conteúdo
- [11.5] da metainformação que lhe está associada.

Ponto 12

Este ponto tem a ver a ver com a existência do documento que defina o Plano de Preservação Digital.

Ponto 13

Este ponto diz respeito à produção de testes preventivos sobre cópias da Informação de Arquivo Eletrónica com o fim de antecipar falhas e permitir reformulações do Plano de Preservação Digital. Estes testes podem incluir a utilização do software e hardware, ou aplicação das estratégias de preservação.

Ponto 14

Este ponto implica a produção de cópias de segurança da Informação de Arquivo Eletrónica a preservar, para evitar quaisquer perdas motivadas por erros causadas durante a implementação.

Ponto 15

Este ponto refere-se à implementação o Plano de Preservação Digital produzido. Aqui pretende-se saber se o PPD já foi implementado.

A Lista de requisitos para a prática de preservação digital

Indicar na coluna *Comentário* a seguinte informação:

Se o requisito for considerado cumprido assinalar com uma x a opção “Sim”

Se o requisito for considerado não cumprido com uma x a opção “Não”

#	Designação	Obrigatoriedade	Comentário	
1	O organismo dispõe de:			
1.1	Plano de classificação	Obrigatório	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
1.2	Tabela de Seleção	Aconselhável	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
2	Foi elaborada uma análise e caracterização da informação eletrónica produzida pela organização no âmbito:			
2.1	do sistema de gestão documental	Obrigatório	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
2.2	das séries documentais [classe de informação ou atividades suportadas]	Obrigatório	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
3	A análise anterior permitiu:			
3.1	definir classes de informação a serem preservadas	Obrigatório	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
3.2	identificar as características dos ficheiros por cada classe de informação definida	Obrigatório	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
3.3	identificar as características técnicas da Informação de Arquivo Eletrónica	Obrigatório	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
3.4	identificar os seguintes fatores inerentes à Informação de Arquivo Eletrónica:			
3.4.1	Usabilidade	Obrigatório	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
3.4.2	Complexidade	Obrigatório	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
3.4.3	Escalabilidade	Obrigatório	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
4	Foi produzida uma avaliação e escolha de:			

#	Designação	Obrigatoriedade	Comentário	
4.1	formatos normalizados de preservação a longo prazo, para cada formato de ficheiro a preservar?	Obrigatório	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
4.2	aplicações informáticas:			
4.2.1	com capacidade de incremento de utilizadores simultâneos	Obrigatório	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
4.2.2	com especificações abertas	Obrigatório	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
4.2.3	que permitam a exportação para formatos normalizados	Obrigatório	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
4.3	soluções armazenamento:			
4.3.1	que permitam o escalamento da capacidade de armazenamento	Obrigatório	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
4.3.2	que incluam sistemas de <i>backup</i>	Obrigatório	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
4.3.3	que incluam sistema de rotinas de controlo e monitorização	Obrigatório	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
5	Analisaram-se e adaptaram-se os esquemas de metainformação:			
5.1	já utilizados no organismo, com elementos de metainformação:			
5.1.1	Descritiva	Obrigatório	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
5.1.2	Técnica	Obrigatório	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
5.1.3	Estrutural	Obrigatório	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
5.1.4	Preservação	Aconselhável	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
5.2	a serem implementados e que contenham elementos de metainformação:			
5.2.1	Descritiva	Obrigatório se aplicável	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
5.2.2	Técnica	Obrigatório se aplicável	Sim Não	<input type="checkbox"/> <input type="checkbox"/>

#	Designação	Obrigatoriedade	Comentário	
5.2.3	Estrutural	Obrigatório se aplicável	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
5.2.4	Preservação	Obrigatório se aplicável	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
6	Foram elaborados e analisados estudos no intuito de identificar os custos de soluções de Preservação na fase de desenvolvimento e implementação em termos de:			
6.1	<i>software</i>	Obrigatório	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
6.2	<i>hardware</i>	Obrigatório	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
6.3	preparação e adaptação de infraestruturas	Obrigatório	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
6.4	recursos humanos	Obrigatório	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
7	Foram analisados e identificados os custos de funcionamento, manutenção e atualização em termos de:			
7.1	parque informático	Obrigatório	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
7.2	infraestrutura	Obrigatório	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
7.3	formação	Obrigatório	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
7.4	recursos humanos	Obrigatório	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
8	Foram afetados os recursos definidos para implementação	Obrigatório se aplicável	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
9	Foi definido um cronograma para execução da implementação	Aconselhável	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
10	Foram definidas medidas de controlo internas a nível de:			
10.1	procedimentos de preservação	Obrigatório	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
10.2	formatos para preservação	Obrigatório	Sim Não	<input type="checkbox"/> <input type="checkbox"/>

#	Designação	Obrigatoriedade	Comentário	
10.3	soluções de software alternativas para um determinado formato	Obrigatório	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
10.4	soluções de armazenamento	Obrigatório	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
10.5	esquemas de metainformação	Obrigatório	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
11	Foram definidas medidas de monitorização interna da Informação de Arquivo Eletrónica preservada, em termos de:			
11.1	integridade	Obrigatório	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
11.2	funcionalidade	Obrigatório	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
11.3	estrutura	Obrigatório	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
11.4	conteúdo	Obrigatório	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
11.5	metainformação associada	Obrigatório	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
12	Foi produzido um documento que defina o PPD de acordo com as soluções propostas anteriormente	Obrigatório	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
13	A solução proposta foi testada preventivamente sobre cópias de Informação de Arquivo Eletrónica, para despistar falhas	Aconselhável	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
14	Foram produzidas cópias de segurança da Informação de Arquivo Eletrónica a preservar?	Aconselhável	Sim Não	<input type="checkbox"/> <input type="checkbox"/>
15	A estratégia de preservação digital foi implementada de acordo com o definido no PPD desenvolvido pelo organismo?	Obrigatório	Sim Não	<input type="checkbox"/> <input type="checkbox"/>



B Folha de Recolha de Dados

O presente anexo contém a Folha de Recolha de Dados e respetivas instruções.

FASE 2

Esta folha - **Identificação** - deve ser preenchida para todos os **SISTEMAS de informação** existentes no organismo, independentemente de estarem ou não em exploração. Por "sistema de informação", no quadro de elaboração do PPD, entende-se um sistema informático que, conceptual e fisicamente, se constitui como unidade autónoma ou autonomizável de outros SI. Os SI têm, frequentemente, expressão informacional em bases de dados ou sistemas de bases de dados relacionadas, mas podem também ser constituído por ficheiros. Quando o SI seja constituído por diferentes subsistemas (módulos ou mesmo aplicações específicas que processam a entrada de dados, armazenamento, recuperação, análise e exploração, reporting, ou outras funções), estes não devem ser considerados individualmente nem para efeitos de identificação do sistema (1) nem para efeitos de avaliação arquivística (2) - só na **FASE 4A**, relativa à caracterização tecnológica, se poderá justificar a decomposição do SI ou, em casos excepcionais, para efeitos de avaliação.

Identificação do sistema de informação	Responsabilidades						
Nome	Administrador (sistema/dados)	Proprietário (sistema/dados)	Utilizador (es) (Organismo(s) ou U.O)	Localização dos dados/informação	Definição formal de responsabilidades individuais	Insourcing	Outsourcing

Registrar, por extenso, a designação oficial do sistema ou, quando não exista, nome pelo qual é conhecido.

Designação do organismo que tem a administração operacional do sistema e dos dados (assegura o armazenamento dos dados, monitorização do sistema, actualização, etc.). Caso exista diferença entre administrador do Sistema e Administrador dos Dados, assinalar com a letra "S" ou "D" respectivamente.

Designação do organismo que tem a propriedade intelectual e/ou responsabilidades de gestão global do sistema e dos dados (decide sobre o desenvolvimento, distribuição, desactivação, etc.). Caso exista diferença entre proprietário do Sistema e o proprietário dos Dados, assinalar com a letra "S" ou "D" respectivamente.
Caso a propriedade do Sistema seja privada (uso condicionado à aquisição de licenças), referir também o organismo que decide sobre a renovação ou não das licenças de utilização.

Nome da(s) unidade(s) orgânica(s) que utilizam o sistema para cumprimento das suas actividades, inserindo dados ou simplesmente utilizando-os para consulta. Quando se trate de um sistema de uso partilhado por vários organismos, se o organismo que está a responder a este inquérito for proprietário ou administrador do sistema deve referir quais os organismos que o partilham. Este campo visa ajudar a identificar os interlocutores para a avaliação arquivística (fase 3B).

Indicar nome do(s) organismo(s) onde os dados residem/estão armazenados.

Campo SIM/NÃO e, em caso afirmativo, indicar qual a expressão dessa formalização de mandatos (ex. despacho oficial, registo no próprio sistema de utilizadores e de permissões...)
Por "responsabilidades individuais" entende-se todo o tipo de permissões necessárias à gestão do sistema (monitorização, registo/alteração de permissões...) e dos dados (criação de dados, consulta, edição, eliminação...).

Campo SIM/NÃO e, em caso afirmativo, indicar a natureza dos serviços prestados (não considerar os contratos de manutenção, a menos que incluam o desenvolvimento do sistema).
Considera-se "insourcing" quando o prestador de serviços é entidade do sector administrativo do Estado.

Campo SIM/NÃO e, em caso afirmativo, indicar a natureza dos serviços prestados (não considerar os contratos de manutenção, a menos que incluam o desenvolvimento do sistema).
Considera-se "outsourcing" quando o prestador de serviços é entidade do sector privado ou do sector público empresarial.

FASE 3 A (Sistemas)						
Esta folha - Avaliação 3A (Sistemas) - deve ser preenchida apenas para os SISTEMAS que não estão referenciados em tabela de selecção em vigor, aprovada nos termos da lei (sistemas ainda não avaliados).				Avaliação arquivística de sistemas de informação		
Cod. Classif.	Nº de Ref.	Nome do sistema	Área de negócio (Função / Subfunção)	Âmbito e conteúdo	Diplomas juridico-administrativos	Utilização do sistema
Código do Plano de Classificação, se aplicável	Nº de identificação sequencial do sistema a avaliar, com o prefixo SI (ex.: SI 1, SI 2... SI 16...)	Designação pela qual o sistema é conhecido (mesma designação usada na folha de identificação do sistema).	Descrição sucinta das actividades, administrativas ou técnicas, e procedimento(s) administrativo, documental ou outro que o sistema serve ou suporta.	Registrar qualquer normativo (lei, decreto-lei, portaria, directiva, regulamento, instrução, despacho ou outro) que determine as características do sistema ou da informação que nele reside, ou que influencie os seus prazos de conservação ou destino final .	Indicar se o sistema está: Activo (em exploração: cumpre objectivo operacional primário) Semi-Activo (usado apenas para consulta esporádica: o SI foi substituído por outro mais actual; dados não foram migrados, ou foram migrados com erros) Inactivo (sem qualquer utilização administrativa. Ex: No Contexto de uma reforma da administração).	
Registrar a(s) área(s) de negócio em que o sistema é utilizado (pode identificar a função que está na base da sua utilização, entendendo se por função qualquer objetivo de alto nível, responsabilidade ou tarefa prescrita como atribuição de uma entidade colectiva pela legislação, política ou mandato). As funções podem ser decompostas em subfunções (ex: gestão de recursos humanos - controlo de assiduidade).						

FASE 3A (Sistemas)

Avaliação arquivística de sistemas de informação

Série/sistema relacionado (nº referência)	Série/sistema relacionado (título)	Série /sistema relacionada (tipo de relação)	Prazo de conservação - efeitos probatórios	Prazo de conservação - efeitos informacionais	Prazo de conservação administrativa	Forma de contagem dos prazos
---	------------------------------------	--	--	---	-------------------------------------	------------------------------

Preencher no caso de haver outros sistemas ou séries que contêm informação relacionada (ver nota ao campo "tipo de relação") com a informação existente no sistema em análise. Indique aqui o **nº de referência atribuído**, nesta mesma folha ou na tabela de selecção, ao sistema/série relacionado. Se a série/sistema relacionado residir noutro organismo, especificar também o organismo em causa.

Se preencheu o campo anterior, indique aqui o **nome do sistema ou título da série relacionada** (que possui informação síntese, duplicada, complementar, antecedente, input ou output relativamente ao sistema em análise).

Período de tempo atribuído à informação que reside no sistema, resultante do seu **valor probatório** (Valor inerente aos documentos de arquivo, na medida em que consignam ou comprovam factos, constituem direitos e obrigações e são reconhecidos como garantia e fundamento de actos, factos e acontecimentos).

Período de tempo (preferencialmente em **anos**) findo o **qual** a informação que reside no sistema não é mais necessária para exploração e/ou consulta pelos utilizadores/interessados primários (produtores).
Caso haja variações de prazos dentro do sistema, especificá-las neste campo e preencher os campos seguintes em conformidade (até ao campo "justificação" do destino final).

O sistema ou série referido nos campos anteriores, possui **informação**... (uma ou mais opções):
S (**síntese** - quando sintetiza o conteúdo informativo do sistema/série em análise).
D (**duplicada** - quando possui, no todo ou em parte, o mesmo conteúdo informativo do sistema/série em análise - não confundir com backups ou réplicas do sistema).
I (**complementar** - quando possui informação adicional que acrescenta significado à informação do sistema/série em análise).
A (**antecedente** - quando se trata de um sistema inactivo ou série fechada, que foi substituído/a pelo sistema/série em análise).
X (**Input** - quando fornece dados ou informação ao sistema em análise. Este tipo de relação pode existir entre dois sistemas ou entre uma série e o sistema (fornecendo a série input ao sistema. Ex: respostas a inquéritos posteriormente inseridas num sistema para processamento estatístico).
O (**Output** - quando a informação, no todo ou em parte, tem origem ou resulta do processamento de dados existentes no sistema em análise. Este tipo de relação pode existir entre dois sistemas ou entre um sistema e uma série (sendo a série um output do sistema. Ex. relatórios produzidos pelo sistema).

Período de tempo atribuído à informação que reside no sistema, resultante do seu **valor informativo** (Valor decorrente da informação veiculada por um documento de arquivo ou outra unidade arquivística para a administração produtora, assim como para outros utilizadores. São especialmente relevantes os que, independentemente do fim para que foram elaborados, testemunham a constituição e funcionamento dessa administração e/ou fornecem dados ou informações sobre pessoas, organizações, locais ou assuntos).

O prazo referido no campo anterior é contado a partir de (indique a opção mais adequada):
- **data de captura dos dados no sistema** (opção preferencial quando a exploração e/ou consulta incide apenas sobre os dados mais recentes, podendo os dados mais antigos ser "descartados");
- **data registada num campo determinado** (indicar qual - ex. data de encerramento de processos);
- **data de desactivação do sistema** (opção preferencial quando a exploração e/ou consulta incide sobre a totalidade dos dados);
- **Outra...** (indique qual).

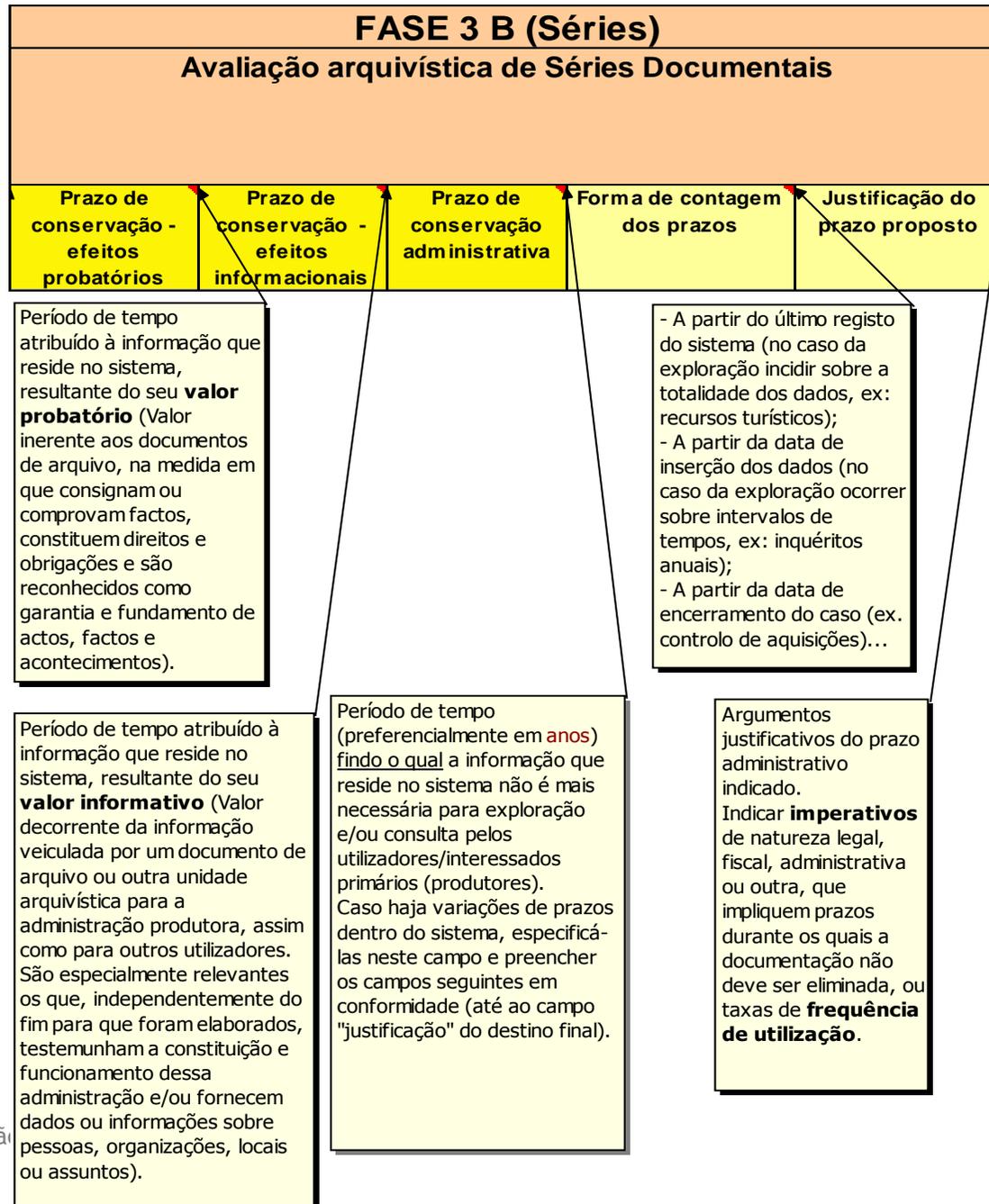
FASE 3A (Sistemas)

Avaliação arquivística de sistemas de informação

Justificação do prazo proposto	Destino Final	Justificação	Notas	Menção de responsabilidade
<p>Referir a natureza do(s) imperativo(s) que justificam o prazo de conservação administrativa proposto (imperativos operacionais, legais, fiscais, etc.), exemplificando a utilização passível de ser observada na fase final do prazo indicado.</p>	<p>DESTINO FINAL - Destino atribuído à informação que reside no sistema, resultante do seu valor probatório e/ou da relevância do seu valor informativo. Registrar: C ou CP ou E C: CONSERVAÇÃO PERMANENTE GLOBAL quando se concluir ter a informação que reside no sistema valor probatório e/ou informativo que impeçam a sua eliminação. Nomeadamente quando o conteúdo informativo não é recuperável noutra sistema, série ou sub série documental de conservação, e revela utilidade para fins de investigação, cumprindo uma função testemunhal da constituição, organização e funcionamento da entidade produtora ou de acontecimentos, pessoas, actos, etc. relevantes para a preservação da memória colectiva. CP: CONSERVAÇÃO PERMANENTE PARCIAL quando se concluir que estamos perante um sistema com conteúdo informativo significativamente homogéneo e repetitivo e cujo interesse informativo não é tão grande que justifique a sua conservação permanente global, mas também se entenda que a eliminação da totalidade da informação é uma opção demasiado radical; Nos casos em que se optar pela conservação permanente parcial, deve ser indicado qual o tipo de informação que se pretende conservar e qual o método de selecção que se pretende implementar (por exemplo, extracção de todos os dados a intervalos de tempo regulares; extracção de dos resultantes de "queries" específicas; etc.) Sublinha-se que a conservação permanente parcial deve ser entendida como um procedimento excepcional, a implementar apenas quando as circunstâncias indiquem uma especial necessidade da mesma e nunca quando haja alternativa viável. E: ELIMINAÇÃO GLOBAL quando: - Se concluir ter o sistema valor probatório e/ou informativo nulo, findo o prazo de conservação administrativa indicado; - Se verificar que a informação contida no sistema é, no todo ou na parte que interessaria conservar, objecto de publicação com registo de depósito legal e se conserve, pelo menos, um exemplar da publicação; - Se verificar que a informação contida no sistema, no todo ou na parte que interessaria conservar, está consubstanciada noutra sistema ou série de conservação permanente no mesmo organismo ou num organismo distinto, desde que ambos pertençam à Administração Pública e o organismo que conserva, a título permanente, a série ou sub série documental, disponha já de uma tabela de selecção publicada em diploma legal, onde a mesma venha expressa.</p>	<p>Argumentos justificativos do destino final indicado. Realce-se que a conservação em função do valor probatório implica requisitos de conservação que assegurem a manutenção desse valor probatório. No caso de se optar por conservação permanente parcial, deve-se também justificar esta opção e a escolha do método de selecção</p>	<p>Este campo é aberto a quaisquer considerações suplementares que se julgue pertinentes para melhor informar o processo de avaliação. Este campo pode igualmente ser utilizado para indicar títulos alternativos para o sistema, este pode ser conhecido pelo nome específico de uma solução ou aplicação informática. Por exemplo: PRIMAVERA, QUIDGEST, SIC, CIME, etc. Nestas situações, pode ser vantajoso indicar esse título, de forma a obter uma identificação mais explícita.</p>	<p>Registrar o nome da pessoa responsável pelo preenchimento da folha de recolha de dados e a data em que o mesmo ocorreu. O seu objectivo é o de facilitar a identificação do responsável pela informação prestada, em caso de dúvida sobre a mesma.</p>

FASE 3 B (Séries)						
Esta folha - <u>Avaliação 3B (Séries)</u> - deve ser preenchida apenas para as SÉRIES que não estão referenciadas em tabela de selecção em vigor, aprovada nos termos da lei, e que no todo ou em parte incluíam documentos electrónicos capturados no Sistema de Gestão de Documentos.				Avaliação arquivística de Séries Documentais		
Cod. Classif.	Nº de Ref.	Nome da série	Função / Subfunção	Âmbito e conteúdo	Diplomas juridico-administrativos	Natureza da utilização
<p>Código da série no Plano de Classificação, se aplicável (ou outro que referencie a série no sistema de gestão de documentos).</p>	<p>Designação da série no Plano de Classificação, se aplicável (ou outro que identifique a série no sistema de gestão de documentos).</p>	<p>Nº de identificação sequencial da série a avaliar, com o prefixo SR (SR 1, SR 2... SR 16...)</p>	<p>Designação da função/subfunção em que a série se inscreve, no Plano de Classificação. Quando não exista Plano de Classificação (ou quando o Plano existente não for funcional), referir função atribuída. Entenda-se por função, qualquer objectivo de alto nível, responsabilidade ou tarefa prescrita como atribuição de uma entidade colectiva pela legislação, política ou mandato. As funções podem ser decompostas em subfunções</p>	<p>Registrar informação para o nível de Função /Subfunção e Série/Subsérie: - Para o nível Função: Fornecer informação narrativa acerca do objectivo da função. - Para o nível de Série: Fornecer informação narrativa acerca das actividades, administrativas ou técnicas, e procedimento(s) administrativo, documental ou outro que se encontrem na sua origem; circuito ou tramitação da documentação ; tipologias documentais (se não constar do título); documentos que enformam a série.</p>	<p>Registrar qualquer normativo (lei, decreto-lei, portaria, directiva, regulamento, instrução, despacho ou outro) que determine a produção documental ou influencie os seus prazos de conservação ou destino final.</p>	<p>Probatória: quando continua a ser necessário manter a admissibilidade legal original; - Informativa: quando apenas é necessário manter o acesso aos dados/ aparência do objecto, mas não a capacidade de processamento original.</p>

FASE 3 B (Séries)				
Avaliação arquivística de Séries Documentais				
Dimensão actual (em múltiplos de byte)	Taxa de crescimento anual (em múltiplos de byte)	Série/sistema relacionado (nº referência)	Série/sistema relacionado (título)	Série relacionada (tipo de relação)
<p>Preencher no caso da série ou sub série em análise conter informação síntese, duplicada, complementar ou output. Registrar o nº de referência da série relacionada, se aplicável.</p>		<p>Registrar o título da série /subsérie que possui informação síntese, duplicada, complementar ou output relativamente série ou sub série em análise. Se esta existir noutro organismo, dar essa informação, especificando, entre parênteses, a designação do mesmo.</p>		
<p>Preencher no caso da série ou sub série em análise conter informação síntese, duplicada, complementar ou output. Registrar: S (Informação síntese - aplica-se se o conteúdo informativo existente numa ou em mais séries documentais é sintetizada ou resumida numa outra série) D (Informação duplicada - verifica-se quando uma série documental contém, no todo ou em parte, o mesmo conteúdo informativo que outra(s) série(s)) I (Informação complementar - aplicável quando o conteúdo informativo de uma série está relacionado com a de outra série acrescentando-lhe informação adicional) A (Série antecedente) O (Output é um tipo de relação que se verifica apenas entre um sistema informático ou base de dados e os respectivos relatórios que são produzidos pelo sistema. Na prática, os outputs são os relatórios (reports) produzidos por um sistema informático. Uma série que seja output de outra tem sempre informação parcialmente duplicada relativamente àquela a partir da qual é produzida. Neste caso, é importante salientar que se o sistema informático que dá origem a esses relatórios (outputs) desaparecer (for destinado a eliminação) os outputs deixarão também de ser produzidos. Esta informação é importante para atribuição de destinos finais.)</p>				



FASE 3 B (Séries)

Avaliação arquivística de Séries Documentais

Destino Final	Justificação	Notas	Menção de responsabilidade
---------------	--------------	-------	----------------------------

Registrar o nome da pessoa responsável pelo preenchimento da folha de recolha de dados e a data em que o mesmo ocorreu. O seu objectivo é o de facilitar a identificação do responsável pela informação prestada, em caso de dúvida sobre a mesma.

Este campo é aberto a quaisquer considerações suplementares por parte da instituição produtora e que se julgue pertinentes para melhor informar o processo de avaliação, podendo ser igualmente usado para apresentar propostas de transferência de suporte, quando for essa a política do organismo em questão. Este campo pode igualmente ser utilizado para indicar títulos alternativos à série documental. No caso das séries em formato electrónico, por vezes, esta pode ser conhecida pelo nome específico de uma solução ou aplicação informática. Por exemplo: PRIMAVERA, QUIDGEST, SIC, CIME, etc. Nestas situações, pode ser vantajoso indicar esse título, de forma a obter uma identificação mais explícita.

Argumentos justificativos do destino final indicado. No caso de se optar por conservação permanente recorrendo a um ou à conjugação de vários métodos de amostragem, deve-se também justificar a opção tomada.

Destino atribuído à série resultante do seu **valor probatório** e/ou da relevância do seu **valor informativo**.
 Registrar: **C** ou **CP** ou **E**
C (CONSERVAÇÃO PERMANENTE GLOBAL quando se concluir ter a série ou sub-série valor probatório e/ou informativo que impeçam a sua eliminação. Nomeadamente que o conteúdo informativo não é recuperável noutras séries ou sub-séries documentais de conservação, revela utilidade para fins de investigação, cumprindo uma função testemunhal da constituição, organização e funcionamento da entidade produtora ou de acontecimentos, pessoas, actos, etc. relevantes para a preservação da memória colectiva.
CP (CONSERVAÇÃO PERMANENTE PARCIAL quando:
 - Se concluir que estamos perante uma série e/ou sub-série muito extensa, com conteúdo informativo significativamente homogéneo e repetitivo e cujo interesse informativo não é tão grande que justifique a sua conservação permanente global, mas também se entenda que a eliminação da totalidade da série é uma opção demasiado radical;
 - Ou se concluir que estamos perante uma série e/ou sub-série constituída por processos ou dossiers cujo interesse informativo é muito desigual.
 Nos casos em que, por estes motivos, se optar pela conservação permanente parcial, deve ser indicado qual o tipo de amostragem que se pretende implementar (exemplar, selectiva, sistemática, aleatória). Porém, esta deve ser entendida como um procedimento excepcional, a implementar apenas quando as circunstâncias indiquem uma especial necessidade do mesmo e nunca quando haja alternativa.
 § N.B. : Algumas entidades produtoras costumam propor a conservação permanente parcial de séries constituídas por diferentes conjuntos documentais com distinto interesse informativo ou probatório, explicitando qual o destino proposto para a generalidade da série, que documentos constituem excepção e porquê.
 Contudo, se a documentação se encontra bem organizada, de modo a identificar as diferentes sub-séries de uma série, deve optar-se pelo preenchimento de uma ficha para cada uma delas, com as respectivas propostas de prazo de conservação e destino final.
 Por exemplo:
 Série: Processos de inspecção
 Sub-série: Relatórios – Conservação permanente
 Sub-série: Documentos de apoio – Eliminação.
E (ELIMINAÇÃO GLOBAL quando:
 - Se concluir ter a série ou sub-série valor probatório e/ou informativo arquivístico nulo;
 - Se verificar que a informação contida na série ou sub-série é objecto de publicação, desde que à publicação tenha sido atribuído um número de depósito legal e no arquivo definitivo se conserve, pelo menos, um exemplar da publicação;
 - Se verificar que os dados essenciais da série ou sub-série estão consubstanciados noutra série ou sub-série de conservação permanente no mesmo organismo ou num organismo distinto, desde que ambos pertençam à Administração Pública e o organismo que conserva, a título permanente, a série ou sub-série documental, disponha já de uma tabela de selecção publicada em diploma legal, onde a mesma venha expressa.

FASE 4A - Caracterização tecnológica de sistemas

Esta folha deve ser preenchida apenas para os sistemas já avaliados, que se encontrem pelo menos numa das situações seguintes:

- Prazo de conservação administrativa superior a 7 anos;
- Destino final de conservação permanente, global ou parcial (independentemente do prazo de conservação administrativa).

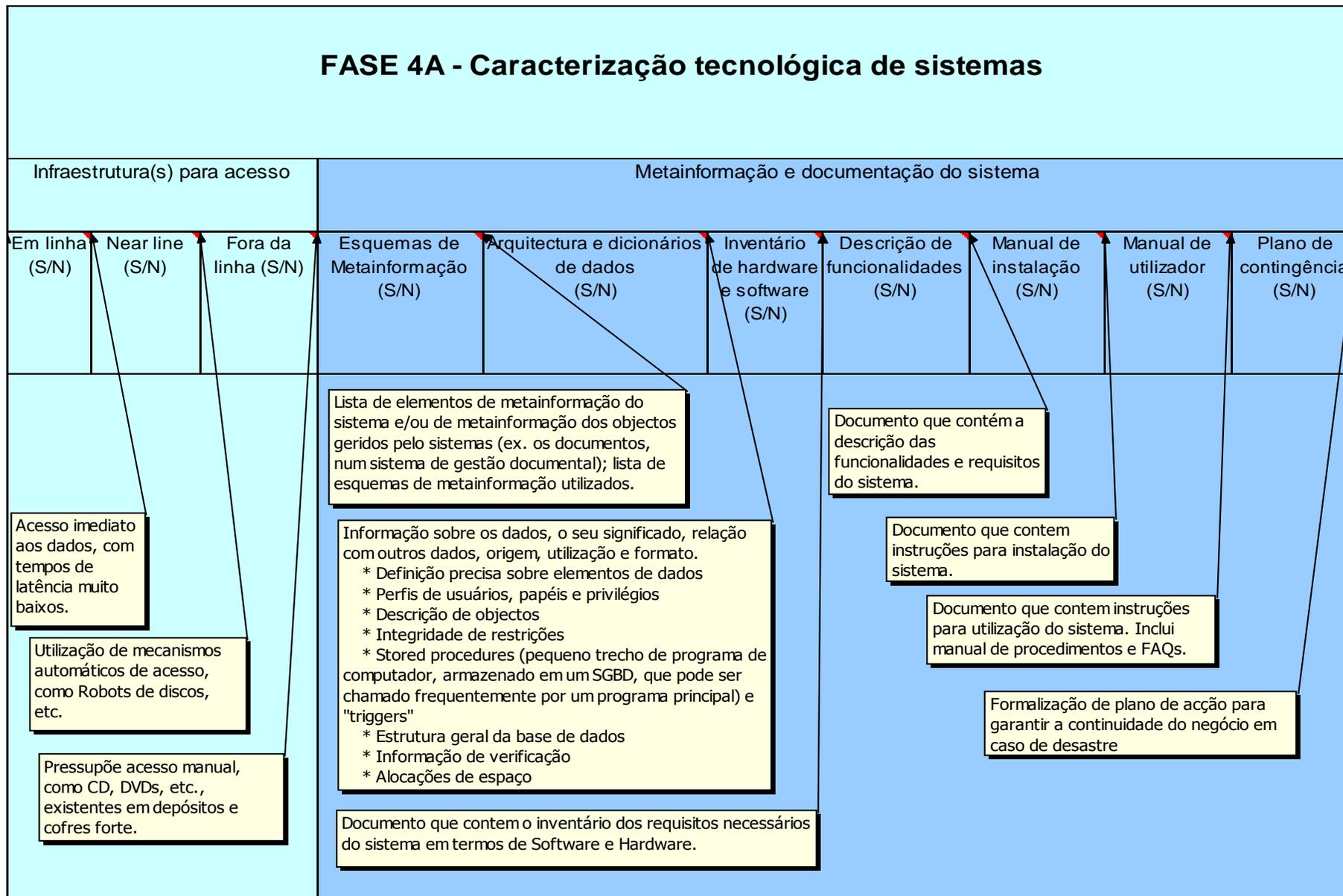
Identificação do sistema			Interoperabilidade do sistema		Crescimento do sistema			
Nome do sistema	Nº. Ref.	Decomposição do sistema (se aplicável)	Nível de dependência do software	Categoria dos dados e formatos utilizados	Modelo de crescimento	Dimensão actual	Crescimento anual	Escalabilidade
<p>Designação pela qual o sistema é conhecido (mesma designação usada na folha de identificação do sistema).</p> <p>Nº de referência extraído de: - folha (Avaliação 3A) não esquecendo o prefixo aí utilizado (ex. SI 1, SI 2...); OU - tabela de selecção aprovada nos termos da lei, com o prefixo TS (ex. TS 1, TS 2... TS 10...) - nesta situação, referenciar também o instrumento legal em causa.</p> <p>Nome dos subsistemas/aplicações que integram o sistema. Este campo só deve ser preenchido quando os elementos de caracterização solicitados apresentem variações entre as diferentes aplicações que integram o SI em análise. Usar uma célula diferente para cada aplicação.</p>			<p>Indicar nível de dependência para com fornecedores privados: - 0 (uso exclusivo de sw não proprietário); - 1 uso de sw proprietário, mas com adequadas facilidades de exportação de dados; - 2 (uso de sw proprietário com especificações fechadas, e sem adequadas facilidades de exportação de dados).</p> <p>Exemplos: dados tabulares: sql, MySQL; dados tabulares: ora, Oracle; dados tabulares: mdb, MAccess 2003; texto estruturado: doc, MWord 2007; texto estruturado: pdf, Acrobat P CS4 imagem: jpeg, Acrobat Ph. CS4; imagem: tiff, Paint (Windows Vista); misto: pdf, Acrobat Prof. CS4; misto: ppt, MPowerpoint 2007; misto: xls, MExcel 2007; texto estruturado: xls, MExcel 2007.</p>		<p>Descrever o modelo. Ex.: - acumulação contínua; - actualização de registos por substituição de dados; - expurgo de registos que perdem o interesse; - outra situação - indique qual.</p> <p>Indique nº de registos e dimensão em múltiplos de bytes.</p> <p>Indique nº de registos e dimensão em múltiplos de bytes.</p> <p>Indicar se o SI tem ou não capacidade de crescimento, sem comprometimento do desempenho, atendendo a: - Capacidade de armazenamento; - Velocidade de pesquisa, visualização /recuperação; - Nº de utilizadores simultâneos.</p>			

FASE 4A - Caracterização tecnológica de sistemas

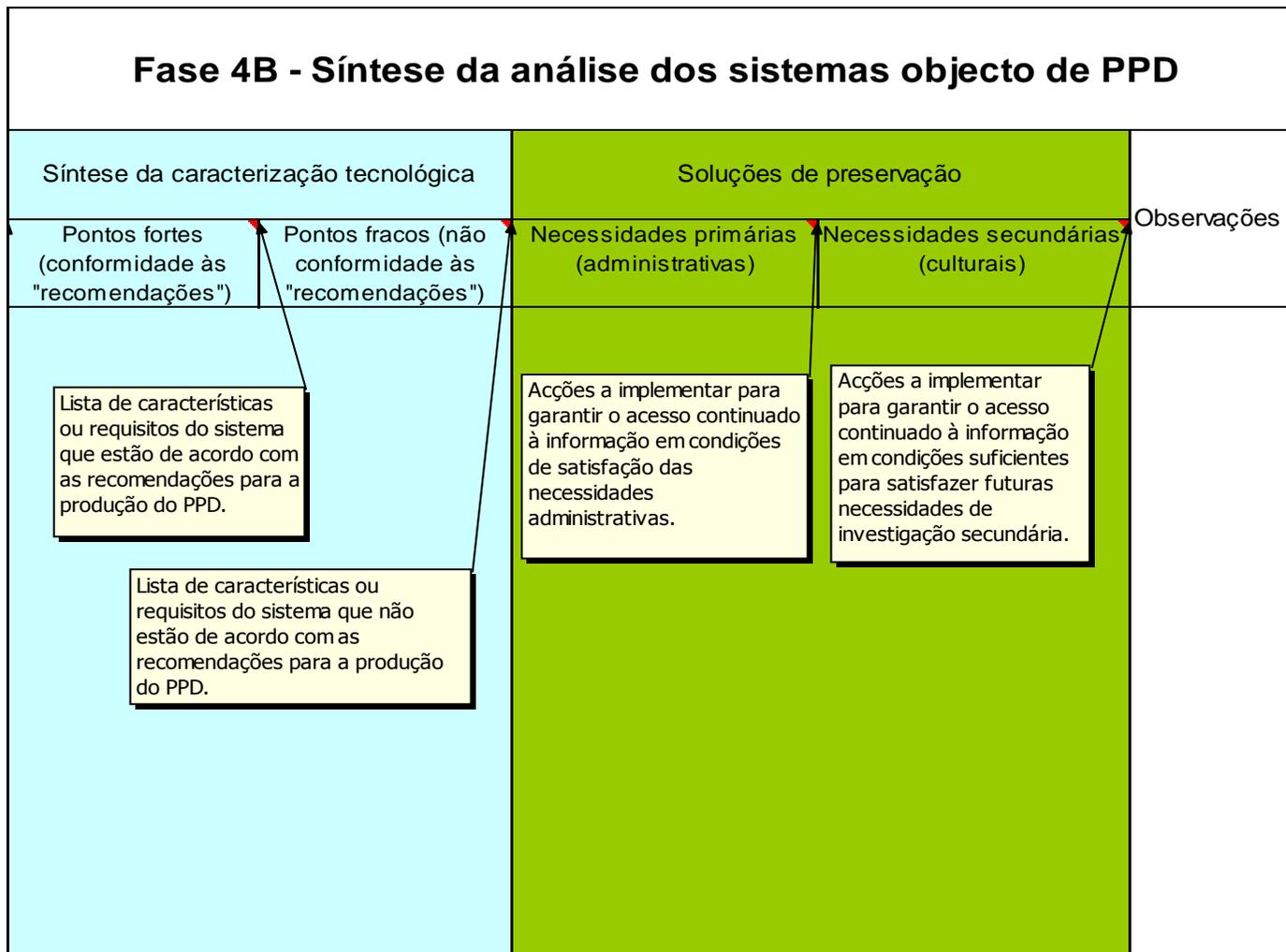
Segurança do sistema (aspectos básicos)							Rotinas de auditoria (registro de ocorrências: quem, o quê, quando)				
Firewall (S/N)	Acesso ao local físico (S/N)	Sistema de detecção de intrusos (S/N)	Password (S/N)	Privilégios de acesso (S/N)	Categorias de segurança dos dados (S/N)	Medidas de recuperação em caso de desastre	Sobre sistema (S/N)	Sobre dados / ficheiros (S/N)	Sobre o armazenamento (S/N)	Produz relatórios (S/N)	Demonstram a fiabilidade do sistema (S/N)
<p>O sistema está ligado a ou contém um dispositivo de uma rede de computadores que tem por objetivo aplicar uma política de segurança a um determinado ponto de controlo da rede.</p> <p>Segurança a nível de acesso ao local onde se encontra o sistema.</p> <p>Existência de Antivirus, Spyware....</p> <p>O acesso ao sistema está limitado através de um sistema de autenticação de utilizadores.</p> <p>Existência de privilégios de acesso diferenciados.</p> <p>Caso o sistema (ou alguns dados do sistema) estejam sujeitos a reservas de comunicabilidade, indicar as categorias de segurança (ex.: muito secreto, secreto, confidencial, reservado)</p>							<p>Registo de ocorrências sobre o sistema?</p> <p>Registo de ocorrências sobre dados/ficheiros?</p> <p>Registo de ocorrências o sistema de armazenamento?</p> <p>Produção de relatórios suficientes a partir das rotinas de auditoria?</p> <p>As rotinas de auditoria (e os seus relatório) permitem verificar todas as ocorrências pertinentes para a demonstração da fiabilidade da informação que reside no sistema?</p>				
<p>Indicar genericamente as medidas previstas para recuperação dos dados, em caso de desastre.</p>											

FASE 4A - Caracterização tecnológica de sistemas

Armazenamento e backup (armazenamento primário (transaccional) + armazenamento secundário (arquivo) + backup (cópias de segurança))							
Servidores	Sistema de armazenamento (NAS, SAN, CAS...)	Computadores individuais (inclui portáteis, PDA...)	Suportes externos portáteis (CD; DVD, Pen, Discos externos...)	Existência de procedimento formais relativos à passagem de dados	Redundância de dados	Modelo de repositório de backups	Frequência de produção de backups
<p>Para armazenamento primário, secundário e/ou backups?</p>					<p>Em caso afirmativo, indicarse se trata de sistema interno (ex. RAID) ou externo (replicação remota).</p>		<p>Indicar se há regularidade e, em caso afirmativo, a periodicidade.</p>
<p>Em caso afirmativo, indicar se é referente à passagem de dados entre sistemas de armazenamento primário e secundário e/ou para a produção de backups.</p>					<ul style="list-style-type: none"> - Não estruturado: armazenado com informações mínimas sobre o que e quando foi armazenado - Total+Incremental: cópia de segurança completa de todos os dados, e posteriormente somente dos dados que foram modificados desde a último backup - Total+Diferencial: Após a cópia de segurança completa ser feita, cada cópia diferencial captura todos os dados criados ou modificados desde a cópia completa, apesar de alguns já poderem ter sido incluídos numa cópia diferencial anterior - Réplica + Incremental invertido (Sincronização): similar ao total+incremental, mas difere na medida que oferece uma cópia que reflecte o estado dos dados da última cópia de segurança e a história reversa das cópias incrementais. - Protecção de dados contínua: o sistema regista imediatamente cada mudança nos dados, o que é geralmente feito diferenças de bytes ou blocos de bytes e não de dados. 		



Identificação do sistema de informação		Síntese da Avaliação Arquivística		
Nome do sistema	Nº de Ref.	Utilização administrativa aos 7 anos	Tolerância a perdas aos 7 anos	Destino final
<p>Designação pela qual o sistema é conhecido (mesma designação usada na folha de identificação do sistema).</p> <p>Nº de referência extraído de: - folha (Avaliação 3A) não esquecendo o prefixo aí utilizado (ex. SI 1, SI 2...); OU - tabela de selecção aprovada nos termos da lei, com o prefixo TS (ex. TS 1, TS 2... TS 10...) - nesta situação, referenciar também o instrumento legal em causa.</p>		<p>7 anos após a produção da informação, a sua utilidade administrativa é:</p> <ul style="list-style-type: none"> - inexistente; - informativa (dispensa funcionalidades do sistema); - exploração (exige que funcionalidades do sistema sejam preservadas). 	<p>7 anos após a produção da informação, a tolerância a perdas é:</p> <ul style="list-style-type: none"> - 0 (a informação continua a ser vital à continuidade do negócio/ cumprimento da missão); - 1 (a informação continua a ser útil para apoio ao ciclo de gestão). - 2 (a informação é apenas utilizada de forma subsidiária ou facultativa). 	<p>Registrar uma das opções:</p> <ul style="list-style-type: none"> - Eliminação - Conservação permanente Parcial - Conservação permanente Total.



Exemplo de preenchimento da Folha de Recolha de Dados

FASE 2							
Identificação do sistema de informação	Responsabilidades						
Nome	Administrador (sistema/dados)	Proprietário (sistema/dados)	Utilizador (es) (Organismo(s) ou U.O)	Localização dos dados/ informação	Definição formal de responsabilidades individuais	Insourcing	Outsourcing
Sistema Electrónico de Gestão de Arquivos	UO Informática	Organismo Empresa produtora	Arquivos dependentes	SI sediado na sede da organização	Sim		Sim (contrato de manutenção)

FASE 3 A (Sistemas)						
Esta folha - <u>Avaliação 3A (Sistemas)</u> - deve ser preenchida apenas para os SISTEMAS que não estão referenciados em tabela de selecção em vigor, aprovada nos termos da lei (sistemas ainda não avaliados).				Avaliação arquivística de sistemas de informação		
Cod. Classif.	Nº de Ref.	Nome do sistema	Área de negócio (Função / Subfunção)	Âmbito e conteúdo	Diplomas juridico-administrativos	Utilização do sistema
1.23.456	SI 002	Sistema Electrónico de Gestão de Arquivos	Gestão Documental	Registo de correspondência de entrada, saída e internos ; arquivo de documentos produzidos nos vários sectores; automatização de procedimentos (workflows)	portaria de gestão documental (aguarda publicação)	Activo

FASE 3A (Sistemas)

Avaliação arquivística de sistemas de informação						
Série/sistema relacionado (nº referência)	Série/sistema relacionado (título)	Série /sistema relacionada (tipo de relação)	Prazo de conservação - efeitos probatórios	Prazo de conservação - efeitos informacionais	Prazo de conservação administrativa	Forma de contagem dos prazos
	Todas as séries documentais com documentos registados no SEGA	O; I; X	1-Aplicação da TS ao repositório 2-Os dados de gestão são conservados permanentemente.	Dependente dos documentos existentes no repositório e respectiva metainformação registada no sistema	1-Aplicação da TS ao repositório 2-Os dados de gestão são conservados permanentemente.	1- Aplicação dos prazos de acordo com o estabelecido na TS à data do encerramento do processo ou à data do último documento da pasta; 2- Data do registo

FASE 3A (Sistemas)

Avaliação arquivística de sistemas de informação				
Justificação do prazo proposto	Destino Final	Justificação	Notas	Menção de responsabilidade
1 - As séries electrónicas/digitalizadas serão mantidas no sistema de gestão documental enquanto estiverem a decorrer os prazos de conservação administrativa definidos na tabela de selecção, findos os quais o sistema de gestão documental aplica o destino final devido. 2 - Por dificuldades de desagregação da informação os dados de gestão são conservados permanentemente. Esta opção deve ser questionada se vier a ser possível estabelecer mecanismos de automatismo que permitam identificar e eliminar/transferir informação de gestão relativa a documentos simples e compostos acompanhando o destino das respectivas agregações.	C	1 - A justificação para a conservação no repositório das séries documentais encontra-se na TS 2 - Os dados de gestão são conservados permanentemente (no sistema original ou em repositório) por assegurarem os elementos de autenticidade e fidedignidade dos documentos em repositório.		António Baptista

FASE 3 B (Séries)						
Esta folha - <u>Avaliação 3B (Séries)</u> - deve ser preenchida apenas para as SÉRIES que não estão referenciadas em tabela de selecção em vigor, aprovada nos termos da lei, e que no todo ou em parte incluam documentos electrónicos capturados no Sistema de Gestão de Documentos.				Avaliação arquivística de Séries Documentais		
Cod. Classif.	Nº de Ref.	Nome da série	Função / Subfunção	Âmbito e conteúdo	Diplomas juridico-administrativos	Natureza da utilização
130.10	SR 11	Registos referentes a cheques devolvidos	Gestão da cobrança	Esta série abrange os registos de cheques devolvidos, por falta de provisão ou outros motivos, cujos dados são: - Código banco, código agência, n.º conta, n.º cheque, valor cheque, data emissão, data devolução, motivo devolução, NIB devolução, valor encargos bancários, data ofício, estado, data estado.	DL n.º 7-B/3102, de 20.01	Informativo

FASE 3 B (Séries)				
Avaliação arquivística de Séries Documentais				
Dimensão actual (em múltiplos de byte)	Taxa de crescimento anual (em múltiplos de byte)	Série/sistema relacionado (nº referência)	Série/sistema relacionado (título)	Série relacionada (tipo de relação)
59Gb	500Mb	Portaria de Gestão Documental (Portaria 132983/3109): 109,	XPTO, YTPO, ZPTO	I - XPTO X - YPTO O - ZPTO

FASE 3 B (Séries)				
Avaliação arquivística de Séries Documentais				
Prazo de conservação - efeitos probatórios	Prazo de conservação - efeitos informacionais	Prazo de conservação administrativa	Forma de contagem dos prazos	Justificação do prazo proposto
2 anos	Registo da regularização em conta corrente da AB/XNE	Suficiente para qualquer esclarecimento	1	E

FASE 3 B (Séries)			
Avaliação arquivística de Séries Documentais			
Destino Final	Justificação	Notas	Menção de responsabilidade
E	Sem valor secundário	Nos subsídios, o prazo conta-se a partir do pagamento	Maria de Matos

FASE 4A - Caracterização tecnológica de sistemas

Esta folha deve ser preenchida apenas para os sistemas já avaliados, que se encontrem pelo menos numa das situações seguintes:

- Prazo de conservação administrativa superior a 7 anos;
- Destino final de conservação permanente, global ou parcial (independentemente do prazo de conservação administrativa).

Identificação do sistema			Interoperabilidade do sistema		Crescimento do sistema			
Nome do sistema	Nº. Ref.	Decomposição do sistema (se aplicável)	Nível de dependência do software	Categoria dos dados e formatos utilizados	Modelo de crescimento	Dimensão actual	Crescimento anual	Escalabilidade
Sistema Electrónico de Gestão de Arquivos	SI 002		1	Dados Tabulares; Sql	acumulação	17 GB	100Mb	Sim

FASE 4A - Caracterização tecnológica de sistemas											
Segurança do sistema (aspectos básicos)							Rotinas de auditoria (registo de ocorrências: quem, o quê, quando)				
Firewall (S/N)	Acesso ao local físico (S/N)	Sistema de detecção de intrusos (S/N)	Passw ord (S/N)	Privilégios de acesso (S/N)	Categorias de segurança dos dados (S/N)	Medidas de recuperação em caso de desastre	Sobre sistema (S/N)	Sobre dados / ficheiros (S/N)	Sobre o armazenamento (S/N)	Produz relatórios (S/N)	Demonstram a fiabilidade do sistema (S/N)
Sim	Sim	Não	Sim	Sim	Sim	Nao	Não	Não	Não	Não	Não

FASE 4A - Caracterização tecnológica de sistemas							
Armazenamento e backup (armazenamento primário (transaccional) + armazenamento secundário (arquivo) + backup (cópias de segurança))							
Servidores	Sistema de armazenamento (NAS, SAN, CAS...)	Computadores individuais (inclui portáteis, PDA ...)	Suportes externos portáteis (CD; DVD, Pen, Discos externos...)	Existência de procedimento formais relativos à passagem de dados	Redundância de dados	Modelo de repositório de backups	Frequência de produção de backups
Primário	SAN	Não	Não	Não	Sim (RAID)	Total	Diário

FASE 4A - Caracterização tecnológica de sistemas									
Infraestrutura(s) para acesso			Metainformação e documentação do sistema						
Em linha (S/N)	Near line (S/N)	Fora da linha (S/N)	Esquemas de Metainformação (S/N)	Arquitectura e dicionários de dados (S/N)	Inventário de hardware e software (S/N)	Descrição de funcionalidades (S/N)	Manual de instalação (S/N)	Manual de utilizador (S/N)	Plano de contingência (S/N)
Sim	Não	Não	Sim	Não	Não	Não	Sim	Sim	Sim

Fase 4B - Síntese da análise dos sistemas objecto de PPD				
Identificação do sistema de informação		Síntese da Avaliação Arquivística		
Nome do sistema	Nº de Ref.	Utilização administrativa aos 7 anos	Tolerância a perdas aos 7 anos	Destino final
Sistema Electrónico de Gestão de Arquivos	SI 002	a) Inexistente ou Exploração, conforme as séries documentais sejam eliminadas antes ou depois do prazo de 7 anos; b) Exploração para a metainformação	a) 2 ou 0 b) 0	CP

Fase 4B - Síntese da análise dos sistemas objecto de PPD				
Síntese da caracterização tecnológica		Soluções de preservação		Observações
Pontos fortes (conformidade às "recomendações")	Pontos fracos (não conformidade às "recomendações")	Necessidades primárias (administrativas)	Necessidades secundárias (culturais)	
<p>a) Permite exportar para XML.</p> <p>b) Independente da plataforma aplicacional</p>	<p>a) Dependência de fabricante externo.</p> <p>b) Não foram analisados e identificados os custos de funcionamento, manutenção e actualização;</p> <p>c) Não foi produzida uma avaliação e escolha de formatos normalizados de preservação a longo prazo para cada formato de ficheiro a preservar;</p> <p>d) Não foram feitos estudos de custos de soluções de preservação;</p> <p>e) Não foram definidas medidas de auditoria internas;</p> <p>f) Não foram definidas medidas de monitorização interna dos documentos de arquivo electrónicos preservados.</p> <p>g) Aplicação proprietária</p> <p>h) Inexistência de:</p> <ul style="list-style-type: none"> - esquemas de metainformação; - documentação do sistema, excepto o manual de utilizador 	<p>a) Monitorização dos dados, com a perspectiva de migração recorrendo à actualização de versões e</p> <p>b) Monitorização dos dados, com a perspectiva de migração para nova versão do sistema ou para sistemas concorrentes</p>	<p>Migração de dados para formatos de preservação</p>	

C Exemplo de um ficheiro PREMIS contendo MI técnica NISO Z39.87

```
<?xml version="1.0" encoding="UTF-8"?>
<premis xmlns="http://www.loc.gov/standards/premis/v1"
  xmlns:mix="http://www.loc.gov/mix/"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.loc.gov/standards/premis/v1
    http://www.loc.gov/standards/premis/v1/PREMIS-v1-1.xsd
    http://www.loc.gov/mix/http://www.loc.gov/mix/mix.xsd">
  <object>
    <objectIdentifier>
      <objectIdentifierType>Custom</objectIdentifierType>
      <objectIdentifierValue>R2006.01.0001</objectIdentifierValue>
    </objectIdentifier>
    <preservationLevel>full</preservationLevel>
    <objectCategory>representation</objectCategory>
    <relationship>
      <relationshipType>structural</relationshipType>
      <relationshipSubType>hasroot</relationshipSubType>
      <relatedObjectIdentification>
        <relatedObjectIdentifierType>Custom</relatedObjectIdentifierType>
        <relatedObjectIdentifierValue>F2006.01.0001</relatedObjectIdentifierValue>
        <relatedObjectSequence>0</relatedObjectSequence>
      </relatedObjectIdentification>
    </relationship>
    <relationship>
      <relationshipType>structural</relationshipType>
      <relationshipSubType>haspart</relationshipSubType>
      <relatedObjectIdentification>
        <relatedObjectIdentifierType>Custom</relatedObjectIdentifierType>
        <relatedObjectIdentifierValue>F2006.01.0001.0000001</relatedObjectIdentifierValue>
        <relatedObjectSequence>1</relatedObjectSequence>
      </relatedObjectIdentification>
    </relationship>
    <linkingIntellectualEntityIdentifier>
      <linkingIntellectualEntityIdentifierType>Custom</linkingIntellectualEntityIdentifierType>
      <linkingIntellectualEntityIdentifierValue>1</linkingIntellectualEntityIdentifierValue>
    </linkingIntellectualEntityIdentifier>
  </object>
  <object>
    <objectIdentifier>
      <objectIdentifierType>Custom</objectIdentifierType>
      <objectIdentifierValue>F2006.01.0001</objectIdentifierValue>
    </objectIdentifier>
    <preservationLevel>full</preservationLevel>
    <objectCategory>File</objectCategory>
    <objectCharacteristics>
      <compositionLevel>1</compositionLevel>
      <size>3156</size>
      <format>
        <formatDesignation>
          <formatName>image/xml</formatName>
        </formatDesignation>
        <formatRegistry>
          <formatRegistryName>JHOVE</formatRegistryName>
          <formatRegistryKey>XML1.0</formatRegistryKey>
        </formatRegistry>
        <formatRegistry>
          <formatRegistryName>MIME</formatRegistryName>
          <formatRegistryKey>text/xml</formatRegistryKey>
        </formatRegistry>
      </format>
    </objectCharacteristics>
    <creatingApplication>
      <creatingApplicationName>MakeDVDImage.pl</creatingApplicationName>
      <dateCreatedByApplication>2006-06-09T15:34:18Z</dateCreatedByApplication>
    </creatingApplication>
    <storage>
      <contentLocation>
        <contentLocationType>URI</contentLocationType>
      </contentLocation>
    </storage>
  </object>
</premis>
```

```

    <contentLocationValue>R2006.01.0001/F2006.01.0001.mets</contentLocationValue>
  </contentLocation>
  <storageMedium>DVD</storageMedium>
</storage>
</object>
<object>
  <objectIdentifier>
    <objectIdentifierType>Custom</objectIdentifierType>
    <objectIdentifierValue>F2006.01.0001.0000001</objectIdentifierValue>
  </objectIdentifier>
  <preservationLevel>full</preservationLevel>
  <objectCategory>File</objectCategory>
  <objectCharacteristics>
    <compositionLevel>0</compositionLevel>
    <fixity>
      <messageDigestAlgorithm>MD5</messageDigestAlgorithm>
      <messageDigest>5d5679af155d9752fab9a42e6c724f0a</messageDigest>
    </fixity>
    <size>482190</size>
    <format>
      <formatDesignation>
        <formatName>image/tiff</formatName>
      </formatDesignation>
      <formatRegistry>
        <formatRegistryName>JHOVE</formatRegistryName>
        <formatRegistryKey>TIFF5.0</formatRegistryKey>
      </formatRegistry>
      <formatRegistry>
        <formatRegistryName>MIME</formatRegistryName>
        <formatRegistryKey>image/tiff</formatRegistryKey>
      </formatRegistry>
    </format>
    <significantProperties>
      <mix:mix>
        <mix:BasicImageParameters>
          <mix:Format>
            <mix:MIMEType>image/tiff</mix:MIMEType>
            <mix:ByteOrder>little-endian</mix:ByteOrder>
            <mix:Compression>
              <mix:CompressionScheme>1</mix:CompressionScheme>
            </mix:Compression>
            <mix:PhotometricInterpretation>
              <mix:ColorSpace>0</mix:ColorSpace>
            </mix:PhotometricInterpretation>
            <mix:Segments>
              <mix:StripOffsets>8</mix:StripOffsets>
              <mix:RowsPerStrip>2339</mix:RowsPerStrip>
              <mix:StripByteCounts>481834</mix:StripByteCounts>
            </mix:Segments>
            <mix:PlanarConfiguration>1</mix:PlanarConfiguration>
          </mix:Format>
          <mix:File>
            <mix:Orientation>1</mix:Orientation>
          </mix:File>
        </mix:BasicImageParameters>
        <mix:ImageCreation>
          <mix:ScanningSystemCapture>
            <mix:ScanningSystemSoftware>
              <mix:ScanningSoftware>KofaxImageProcessingPlatformv1.10.018;KIPPTIFFStorageFilterv110.018</mix:ScanningSoftware>
            </mix:ScanningSystemSoftware>
          </mix:ScanningSystemCapture>
          <mix:DateTimeCreated>1992-05-09T15:10:57</mix:DateTimeCreated>
        </mix:ImageCreation>
        <mix:ImagingPerformanceAssessment>
          <mix:SpatialMetrics>
            <mix:SamplingFrequencyUnit>2</mix:SamplingFrequencyUnit>
            <mix:XSamplingFrequency>200</mix:XSamplingFrequency>
            <mix:YSamplingFrequency>200</mix:YSamplingFrequency>
            <mix:ImageWidth>1648</mix:ImageWidth>
            <mix:ImageLength>2339</mix:ImageLength>
          </mix:SpatialMetrics>
          <mix:Energetics>
            <mix:BitsPerSample>1</mix:BitsPerSample>
          </mix:Energetics>
        </mix:ImagingPerformanceAssessment>
      </mix:mix>
    </significantProperties>
  </objectCharacteristics>
</object>

```

```
<mix:SamplesPerPixel>1</mix:SamplesPerPixel>
</mix:Energetics>
</mix:ImagingPerformanceAssessment>
</mix:mix>
</significantProperties>
</objectCharacteristics>
<creatingApplication>
  <creatingApplicationName>tiffcp(libtiff-tools)</creatingApplicationName>
  <creatingApplicationVersion>3.7.3-1ubuntu1.1</creatingApplicationVersion>
  <dateCreatedByApplication>2006-06-09T15:34:15Z</dateCreatedByApplication>
</creatingApplication>
<originalName>3</originalName>
<storage>
  <contentLocation>
    <contentLocationType>URI</contentLocationType>
    <contentLocationValue>R2006.01.0001/F2006.01.0001.0000001.tiff
  </contentLocationValue>
  </contentLocation>
  <storageMedium>DVD</storageMedium>
</storage>
</object>
</premis>
```

D Recursos Aconselhados

ALVES, Ivone [et. al.] - *Dicionário de terminologia arquivística*. Lisboa: IBL, 1993.

ADOBE - *Tiff developer information site*. 2002 URL: [Consult. 10 Dez. 2007]
<<http://partners.adobe.com/public/developer/tiff>>.

ANDERSSON, Ulf - *Short Version of the Sesam Report: Philosophy and Rules Concerning Electronic Archives and Authenticity*. Proceedings of the DLM-Forum on Electronic Records. Brussels, 18-20 December 1996.

ASSOCIAÇÃO PARA O DESENVOLVIMENTO DA SOCIEDADE DA INFORMAÇÃO - Glossário da Sociedade da Informação - Versão 2007

BARBEDO, Francisco *et al.* - *RODA: Relatório Final*, Março 2007, Direcção Geral de Arquivos e Universidade do Minho, [Consult. 10 Dez. 2007] URL:<
<http://roda.iantt.pt/pt/system/files/roda-relatorioFinal.pdf>>;

BARBEDO, Francisco *et al.* - *RODA: Repositório de Objetos Digitais Autênticos*, in CONGRESSO NACIONAL DE BIBLIOTECÁRIOS, ARQUIVISTAS E DOCUMENTALISTAS, 9, Ponta Delgada, Portugal, 2007 - "Bibliotecas e arquivos : informação para a cidadania, o desenvolvimento e a inovação: actas" [CD-ROM]. [S.l.] : Associação Portuguesa de Bibliotecários, Arquivistas e Documentalistas, 2007. ISBN 978-972-9067-37-2.

BARBEDO, Francisco - *RODA - De projeto a programa estratégico para a Preservação Digital*, in *Boletim Arquivos Nacionais*, nº19, Janeiro-Março 2007, Lisboa. ISSN1645-5460;

BARBEDO, Francisco; GOMES, Eugénia; HENRIQUES, Cecília - *SIADE caderno 1: Contexto de Suporte* [Em linha]. Lisboa: Instituto dos Arquivos Nacionais; Instituto de Informática, 2000 [Consult. 10 Dez. 2007] Disponível em:
<http://www.iantt.pt/downloads/SIADE_Caderno1.pdf>.ISBN 972-8107-59-5

BARBEDO, Francisco; GOMES, Eugénia; HENRIQUES, Cecília; LAGOAS, José - *SIADE caderno 2: Recomendações para a Gestão de Documentos de Arquivo Eletrónicos (MOREQ)*. Lisboa: Instituto dos Arquivos Nacionais; Instituto de Informática, 2000 [Consult. 10 Dez. 2007] Disponível em: <URL:http://www.iantt.pt/downloads/SIADE_Caderno2.pdf>. ISBN 972-8107-59-5

BEARMAN, David - *Electronic Evidence. Strategies for Managing Records in Contemporary Organizations*. Pittsburgh: Archives & Museum Informatics. 1994. ISBN: 1-885626-08-8

BEARMAN, David - *Reality and Chimeras in the Preservation of Electronic Records*, in *D-Lib Magazine*, Abril 1999, disponível em: <www.dlib.org/dlib/april99/bearman/04bearman.html>

BEARMAN, David - *Toward a Reference Model for Business Acceptable Communication*, 1994, disponível em: < www.lis.pitt.edu/~nhprc/prog6-5.html>

BESSER, Howard; TRANT Jennifer - *Introduction to Imaging: Issues in Constructing an Image Database*. Santa Monica, CA: Getty Art History Information Program, 1995. ISBN-13: 978-0892363612

BRITISH STANDARDS INSTITUTION - *BS DISC PD0008 A code of Practice for Legal Admissibility and Evidential Weight of Information Stored Electronically*, 1999, BSI, ISBN 10: 0 580 33006 0;

BRITISH STANDARDS INSTITUTION - *BS PD009 Compliance Workbook for use with PD008*. BSI 1996

CONSELHO INTERNACIONAL DE ARQUIVOS - *Documentos de Arquivo Eletrónicos: Manual para arquivistas*. (Estudos, nº 16) trad e adapt. IAN/TT. 2005. [Em linha]. Lisboa: IAN/TT, 2005. [Referência de 8 de Maio de 2006]. [Consult. 10 Dez. 2007] Disponível na internet em: <http://www.ica.org/sites/default/files/ICAEstudo16_PT_4.pdf>.

CONSELHO INTERNACIONAL DE ARQUIVOS - *ISAD(G): Norma geral internacional de descrição arquivística: adoptada pelo Comité de Normas de Descrição, Estocolmo: Suécia, 19-22 de Setembro de 1999*. Trad. Grupo de Trabalho para a Normalização da Descrição em Arquivo. 2ª ed. [Em linha]. Lisboa: IAN/TT, 2004. [Consult. 10 Dez. 2007]. Disponível na internet em: <<http://www.iantt.pt/downloads/ISADG.pdf>>.

CONSELHO INTERNACIONAL DE ARQUIVOS - *ISAAR (CPF): Norma internacional de registo de autoridade arquivística para pessoas colectivas, pessoas singulares e famílias*. Trad. Grupo de Trabalho para a Normalização da Descrição em Arquivo. 2ª ed. [Em linha]. Lisboa: IAN/TT, 2004. [Consult. 10 Dez. 2007]. Disponível na internet em: <<http://www.iantt.pt/downloads/ISAAR2.pdf>>

CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS, *Reference Model for an Open Archival Information System (OAIS) - Blue Book*. Washington: National Aeronautics and Space Administration, 2002.

COOK, Terry - *Electronic Records, Paper Minds: The Revolution in Information Management and Archives in the Post-Custodial and Post-Modernist Era*. Archives and Manuscripts 22 (Novembro 1994): p.300-328.

DECRETO-LEI Nº 447/88. de 10 de Dezembro - Regula a Regula a pré-arquivagem de documentação.

DICIONÁRIO PRÁTICO DE INFORMÁTICA. Alfragide: Editora McGraw-Hill de Portugal, 2000.

DUFF, Wendy - *Will Metadata Replace Archival Description: A Commentary*. Archivaria 39 (Março 1995): 33-38.

FERREIRA, Miguel, BAPTISTA A. A. e RAMALHO J. C. - "An intelligent decision support system for digital preservation," *International Journal on Digital Libraries*, vol. 6, no. 4, 2007

FERREIRA, Miguel - *Introdução à preservação digital - Conceitos, estratégias e actuais consensos*. Guimarães: Escola de Engenharia da Universidade do Minho, 2006 ISBN-13: 978-972-8692-30-8 [em linha] [Consult. 10 Dez. 2007] Disponível na Internet <URL:<https://repositorium.sdum.uminho.pt/handle/1822/5820>>

HENDLEY, T. - *Comparison of Methods & Costs of Digital Preservation*, British Library Research and Innovation Center, West Yorkshire 106, 1998.

HENRIQUES, M., LIBREOTTO, G., RAMALHO, J., e HENRIQUES, P. - *Bidirectional conversion between xml documents and relational data bases*. International conference on CSCW in design, 2002. Disponível na Internet <URL: <http://hdl.handle.net/1822/601>>

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. ISO/TR 15801: *Electronic Imaging-Information stored electronically - Recommendations for trustworthiness and reliability*. Genebra: ISO, 2004.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. ISO/IEC 17799: *Information technology -- Security techniques -- Code of practice for information security management*. Genebra: ISO, 2005.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. ISO/TR 18492: *Long-term preservation of electronic document-based information*. Genebra: ISO, 2005.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. ISO 18934: *Imaging materials - Multiple media archives - Storage environment*. Genebra: ISO, 2006.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. ISO 23081-1: *Information and documentation - Records management processes - Metadata for records - Part 1: Principles*. Genebra: ISO, 2006.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. ISO/DIS 23081-2: *Information and documentation - Records management processes - Metadata for records - Part 2: Conceptual and implementation issues*. Genebra: ISO, 2007.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. ISO/IEC 27001: *Information technology -- Security techniques -- Information security management systems - Requirements*. Genebra: ISO, 2005.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. ISO/IEC 27002: *Information technology -- Security techniques -- Code of practice for information security management*. Genebra: ISO, 2005.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. ISO/IEC 27006: *Information technology -- Security techniques -- Requirements for bodies providing audit and certification of information security management systems*. Genebra: ISO, 2007.

INSTITUTO PORTUGUÊS DA QUALIDADE - NP 4438-1, *Informação e documentação. Gestão de documentos de arquivo. Parte 1: Princípios directores*. Lisboa: Instituto Português da Qualidade, 2005.

INSTITUTO PORTUGUÊS DA QUALIDADE - NP 4438-2, *Informação e documentação. Gestão de documentos de arquivo. Parte 2: Recomendações de aplicação*. Lisboa: Instituto Português da Qualidade, 2005.

INSTITUTO DOS ARQUIVOS NACIONAIS/TORRE DO TOMBO - *Manual para a gestão de documentos*. Coord. Madalena Garcia e Maria João Pires de Lima; elab. Cecília Henriques, Francisco Barbedo e Luís Montalvão. 2ª ed. Lisboa: IAN/TT, 2004.

LAVOIE, B. F. - *The open archival information system reference model: Introductory guide*. Digital Preservation Coalition. 2004

MACNEIL, Heather - *Metadata Strategies and Archival Description: Comparing Apples to Oranges*. *Archivaria* 39 (Março 1995): 22-32.

MORELLI, Jeffrey - *Defining Electronic Records: a Terminology Problem... or Something More*. *Electronic Information Resources and Historians: European Perspectives*, Eds. Seamus Ross, and Edward Higgs, 83-91. St Katharinen: Scripta Mercaturae Verlag, 1993.

NISO Z39.87-2006, *Data Dictionary - Technical Metadata for Digital Still Images*. Bethesda: National Information Standards Institute; 2006

OCLC e RLG - *PREMIS: Data Dictionary for Preservation Metadata.*, 2005

PARER, Dagmar; PARROTT, Keith - *Management Practices in the Electronic Records Environment*. *Archives and Manuscripts* 22, no. 1 (Maio 1994):106-122.

RLG EAD Advisory Group - *RLG Best Practice Guidelines for Encoded Archival Description*, 2002

ROTHENBERG, J. - *Avoiding Technological Quicksand: Finding a Viable Technical Foundation for Digital Preservation A Report to the Council on Library & Information Resources (CLIR)*, 1999, ISBN - 1-887334-63-7 (disponível em www.clir.org/pubs/reports/rothenberg/contents.html ou www.clir.org/pubs/reports/rothenberg/pub77.pdf)

ROTHENBERG, Jeff - *Ensuring the Longevity of Digital Documents*. *Scientific American*. (Janeiro 1995), 42-47.

SARAMAGO, M. - *Metadados para a preservação digital e aplicação do modelo OAIS*. In VIII Congresso da BAD. 2004.

STEPHENS, David O. - *Digital Preservation in the United Kingdom in THE INFORMATION MANAGEMENT*, vol. 34, nº. 4, Outubro 2000

STEPHENS, David O.; WALLACE, Roderick C. - *Electronic Records Retention: Fourteen Basic Principles in THE INFORMATION MANAGEMENT*, vol. 34, nº. 4, Outubro 2000

WALLACE, David - *Managing the Present: Metadata as Archival Description*. *Archivaria* 39 (Março 1995): 11-21.

WALLACE, David - *Metadata and the Archival Management of Electronic Records: A Review*. *Archivaria* 36 (1993): 87-110.

Endereços eletrónicos

CASPAR Project - <http://www.casparpreserves.eu/caspar-project>

Conselho Internacional de Arquivos - <http://www.ica.org/>

[Conservation on Line](#) (vários links) - <http://palimpsest.stanford.edu/>

DELOS - www.delos.info/

DGARQ - <http://www.dgarq.gov.pt>

Digitale Duurzaamheid - <http://www.digitaleduurzaamheid.nl/home.cfm>

DLM-Forum - http://www.europa.eu.int/historical_archives/dlm_forum

Electronic recordkeeping resources (Gestão de documentos eletrónicos) - <http://www.kshs.org/archives/ermlinks/ermlinks.htm>

Erpanet - <http://www.erpanet.org/>

Estudos de caso (Universidade de Indiana) - <http://www.indiana.edu/~libarche/index.html>

Interpares - www.interpares.org/

Jeff Rothenberg's Digital Longevity Papers - <http://www.panix.com/~jeffr/Prof/digilong.html>

Joint Interoperability Test Command Records Management Application (RMA) - <http://jitc.fhu.disa.mil/recmgt/>

MIP - http://www.dgarq.gov.pt/downloads/MIIP_v1.0a.doc

Model Requirements for the Management of Electronic Records (versão original) - <http://www.cornwell.co.uk/moreq>

National Archives of Australia - <http://www.naa.gov.au/>

National Archives and Records Administration (USA) - <http://www.nara.gov/>

New South Wales ELECTRONIC RECORDKEEPING WEB SITE GUIDE - <http://www.records.nsw.gov.au/>

PDF/A Competence Center - The pdf/a competence center webpage, 2007, URL:<<http://www.pdfa.org>>.

PLANETS - www.planets-project.eu/

Projeto Interpares - <http://www.interpares.org/index.htm>

Projeto Duranti - <http://www.interpares.org/UBCProject/index.htm>

Public Record Office - <http://www.pro.gov.uk/>

Records Continuum Research Group (Meta-informação) - <http://rcrg.dstc.edu.au/>

Research Libraries Group (preservação de documentos eletrónicos) - <http://www.rlg.org/>

RLG - *RLG EAD report card*. 2002, - <http://www.rlg.org/eadreport-card>.

RODA - <http://roda.iantt.pt/>

SOCIETY OF AMERICAN ARCHIVISTS - *Society of American Archivists home page*. 2003 - <http://www.archivists.org>.

SOCIETY OF AMERICAN ARCHIVISTS - *EAD tools survey*. 2006, URL: <http://www.archivists.org/saagroups/ead>.

THE LIBRARY OF CONGRESS. *NISO metadata for images in XML schema official web site*. 2004. - <http://www.loc.gov/standards/mix>

THE LIBRARY OF CONGRESS - *Página oficial do EAD versão de 2002*. - <http://www.loc.gov/ead/>.

THE LIBRARY OF CONGRESS. *Página oficial do METS*. 2006. - <http://www.loc.gov/standards/mets>.

THE LIBRARY OF CONGRESS - *Página oficial do PREMIS*. 2006, - <http://www.loc.gov/standards/premis>.

THE LIBRARY OF CONGRESS - *Sustainability of Digital Formats_Planning for Library of Congress Collection: TIFF rev6.0*, Março 2007. - <http://www.digitalpreservation.gov/formats/fdd/fdd000022.shtml>

E Glossário

API - *Application Programming Interface* (ou **Interface de Programação de Aplicativos**) é um conjunto de rotinas e normas/regras estabelecidos por um software para utilização das suas funcionalidades por programas aplicativos - isto é: programas que não se querem envolver em detalhes da implementação do software, mas apenas utilizar os seus serviços. De modo geral, a API é composta por uma série de funções acessíveis somente por programação, e que permitem utilizar características do software menos evidentes ao utilizador tradicional.

(fonte: Wikipédia Consult. Dez. 2007)

Aplicação - Utilização de técnicas informáticas para resolver um problema de aplicação, que conduz normalmente ao desenvolvimento ou à utilização de programas de aplicação como, por exemplo, folhas de cálculo ou processadores de texto.

Nota: O termo “aplicação informática” é muitas vezes tomado no sentido de “programa de aplicação”.

(fonte: Glossário da sociedade da informação - Versão 2007)

Programa concebido como ferramenta auxiliar para o desempenho de uma determinada tarefa, como processamento de texto, a contabilidade ou a gestão de inventários

(fonte: Dicionário prático de informática - McGraw Hill)

Administrador - Função responsável pela execução rotineira da política de gestão de documentos de arquivos numa organização

(fonte: MoReq)

Arquivo Digital - Consiste numa organização de pessoas e sistemas, que aceitou a responsabilidade de preservar informação e torná-la disponível para uma dada comunidade.

(fonte: OAIS adapt.)

Autenticidade - Um documento de arquivo autêntico é aquele do qual se pode provar que a) ser aquilo que pretende ser, b) ter sido produzido ou enviado pelo alegado produtor ou remetente, e c) ter sido produzido ou enviado no alegado momento de produção ou envio. Para assegurar a autenticidade dos documentos de arquivo, as organizações devem implementar e documentar políticas e procedimentos de controlo da produção, receção, transmissão, manutenção e destino dos mesmos, assegurando que quem produz documentos está devidamente autorizado e identificado e que os documentos estão protegidos contra qualquer ação não autorizada (adicionar, apagar ou alterar informação, uso não autorizado ou imposição de restrições de acesso indevidas).

(fonte: NP 4438-1, p. 14.)

Backup - Cópia de um programa, disco ou dados, destinada a ser arquivada ou a salvaguardar ficheiros valiosos, caso os originais sejam destruídos

(fonte *Dicionário prático de informática* - McGraw Hill)

Backup Não estruturado - Este modelo pode ser constituído apenas por um conjunto de suportes ópticos graváveis, com a informação mínima sobre o que foi replicado e quando. Trata-se do modelo mais fácil de implementar mas, provavelmente, o que piores resultados obterá em termos de recuperação de alto nível.

Backup de Proteção de dados contínua - Neste modelo são registadas todas as alterações do sistema, em vez de calendarizar cópias de segurança periódicas. Esta operação é geralmente realizada através da gravação das diferenças a nível de bytes ou blocos de dados, e não das diferenças a nível de ficheiros. Este modelo difere da simples replicação de discos, uma vez que permite retornar ao estado anterior do registo e, assim, restaurar a anterior imagem de dados.

Backup Réplica + Incremental invertido (Sincronização) Este modelo é similar ao backup Total + Incremental, diferindo no facto de fornecer uma réplica que reflete o estado do sistema, aquando da última cópia de segurança, bem como uma lista das incrementações invertidas.

Um dos benefícios reside no facto de requer uma cópia de segurança total. Cada cópia de segurança incremental é imediatamente aplicada à réplica, e os OD a substituir são movidos para uma incrementação invertida, i.e., é feita uma sincronização. Este modelo não deve utilizar suportes móveis, visto que cada cópia de segurança necessita de ser efetuada por comparação com o *mirror*.

Backup Total + Diferencial - Difere do modelo Total + Incremental no sentido em que, após a cópia de segurança total, cada cópia de segurança parcial deve capturar todos os OD criados ou modificados após aquela, mesmo que já tenham sido incluídos na cópia de segurança parcial anterior. A vantagem é que o restauro apenas obriga à recuperação da cópia de segurança total e a última cópia de segurança parcial.

Backup Total+Incremental - Neste caso pretende-se armazenar várias réplicas dos OD originais da seguinte forma:

1. Em primeiro lugar é feita uma cópia de segurança de todos os OD;
2. Posteriormente poderá ser feito uma cópia de segurança incremental;
3. Finalmente, quando se proceder ao restauro do sistema na totalidade e até determinada data, é necessária a localização da cópia de segurança total, bem como do ponto (data) de restauro pretendido.

Este modelo oferece um alto nível de segurança, em que qualquer OD pode ser restaurado e pode ser utilizado com suportes removíveis, tais como tapes e suportes óticos. O aspeto menos positivo é a existência de longas séries de incrementações e necessidade de espaço.

Classificação: identificação sistemática e organização de atividades e/ou documentos de arquivo em classes, de acordo com conversões logicamente estruturadas, métodos e regras representados num sistema de classificação.

(fonte: NP 4438-1, p. 9)

Código fonte - conjunto de palavras escritas de forma ordenada, contendo instruções numa das linguagens de programação existentes no mercado, de maneira lógica. Entende-se como o código base de um programa de computador, implementando o núcleo de funcionalidade do programa.

Complexidade - quando um documento integra diversas componentes digitais (por exemplo as paginas Web)

Conservação a longo prazo: Ato de manter a informação numa forma independente e acessível a longo prazo

(fonte: OAIS adapt.)

Consumidor - papel das pessoas ou sistemas clientes, que interagem com serviços do arquivo digital, com o fim de pesquisar informação preservada de interesse e acederem a essa informação de forma detalhada. Isto poderá incluir pessoas ou sistemas de arquivos digitais internos ou de outros arquivos digitais.

Cópias de Segurança - ver *Backup*

Cópia de segurança incremental - Cópias de seguranças de dados armazenados e/ou alterados posteriormente à cópia de segurança inicial.

Documento de arquivo Eletrónicos (DAE) - Documento de arquivo produzido, transmitido e conservado, em forma eletrónica

Documento Eletrónicos (DE) - Documento produzido ou recebido, e fixo para armazenamento e manipulação num sistema informático, e que requer a utilização do sistema para ser inteligível pelos humanos

(fonte *Interpares 2, glossary adapt.*)

Documentos estruturados - contêm informação de forma a ser interpretada primariamente por aplicações informáticas (Ex: registos de sistemas contabilísticos, registos de dados de controlo de tráfego aéreo, etc.)

Documentos não estruturados - apresentam a informação de forma a ser inteligível primariamente por humanos (Ex: cartas, memorandos, mensagens de correio Eletrónicos, imagens, fotocópias, imagens digitalizadas, registos áudio e/ou vídeo.

Documento Vital - que asseguram a continuidade do negócio em caso de acidente ou catástrofe

(fonte: NP 4438-2 p.21)

Documento de que a instituição depende em absoluto para a prossecução das suas funções. Ou seja, caso se verificasse um acidente que destruísse esses documentos, o organismo ver-se-ia impedido de prosseguir com eficácia a sua atividade. Em ambiente Eletrónicos, toda a documentação técnica relativa às características do sistema de informação instalado e ao seu desempenho, assim como toda a metainformação descritiva na tramitação documental, terá de ser considerada vital, pois caso se verifique quebra do sistema ela terá papel fundamental para que se proceda ao processo de recuperação.

(fonte: Manual de Gestão de Documentos)

Em linha (Online) - Normalmente o tipo de armazenamento de dados mais acessível, que pode iniciar o restauro em milissegundos. Um exemplo disso é um disco rígido interno ou um conjunto de discos rígidos (ligados a uma SAN). Este tipo de armazenamento é muito conveniente e rápido, mas relativamente caro. Por outro lado, sofre de alguma vulnerabilidade, pois os dados podem ser apagados ou alterados, seja por acidente ou por vírus que elimine dados.

Embebido - Característica de algo com identidade própria (que pode existir autonomamente), que faz parte ou está incluído dentro de um conjunto, que também pode ser uma entidade. Ex: um OD que está embebido noutro OD (imagem embebida num documento de texto)

Emulação - Reprodução do comportamento e resultados de um material digital obsoleto através do desenvolvimento de novo hardware e/ou software para permitir a execução da antiga aplicação de software em computadores no futuro

(fonte Inter pares 2, glossary)

Emulador - Software capaz de reproduzir o comportamento de uma plataforma de hardware e/ou software numa outra que à partida seria incompatível.

(fonte: Manual de Preservação)

Encapsulamento - Preservar, juntamente com o objeto digital, a informação necessária e suficiente para permitir o futuro desenvolvimento de conversores, visualizadores ou emuladores. Esta informação poderá consistir, por exemplo, numa descrição formal e detalhada do formato do objeto preservado.

(fonte: Introdução à preservação digital)

Encriptação - conversão de dados para um código secreto (ou de texto linear em texto cifrado) para transmissão numa rede pública

(fonte Inter pares 2, glossary, adapt.)

Processo de execução de uma transformação complexa de um objeto digital de modo a que este não possa ser apresentado por uma aplicação de uma forma legível ou inteligível, a não ser que a transformação correspondente de descryptografia seja aplicada.

(fonte: *Moreq*)

Escalabilidade - característica do sistema ou equipamento que pode crescer em escala, isto é, que possibilita incrementos de capacidade ou funcionalidades acompanhando as necessidades dos usuários (aumento da capacidade de armazenamento do sistema através da substituição ou instalação de novos discos rígidos; versões mais actuais do formato PDF que permitem a inclusão de informação de dados CAD).

(fonte: *Wikipédia Consult. Dez. 2007*)

Ethernet - Arquitetura de redes locais que suporta um tráfego de alto débito. Nota: A Ethernet é um padrão usual de rede local.

(fonte: *Wikipédia Consult. Dez. 2007*)

Ficheiro - Coleção completa de informações com uma designação, como seja um programa, um conjunto de dados utilizados por um programa ou um documento criado pelo utilizador. O ficheiro constitui a unidade de armazenamento básica que permite ao computador distinguir um conjunto de informações de outro. É uma unidade coerente que o utilizador pode obter, alterar, eliminar, guardar ou enviar para um dispositivo de saída

(fonte *Dicionário prático de informática - McGraw Hill*)

Ficheiro linear - ficheiros cuja informação é passível de ser transferida integralmente para formato analógico.

Fora de Linha (*Off-line*) - Similar ao perto de linha, embora necessite de intervenção humana para disponibilizar as unidades de suporte. Isto pode incluir simplesmente o armazenamento das tapes de cópias de segurança nas estantes. O tempo de acesso deste suporte é superior a uma hora.

Formato - Em geral, designa a estrutura ou aspeto de uma unidade de dados. 2) Disposição de dados num ficheiro que normalmente permite que o documento seja lido ou alterado por uma determinada aplicação.

(fonte: *Dicionário prático de informática - McGraw Hill*)

Organização sequencial de dados em termos dos seus componentes.

(fonte: *OAIS*)

Formato proprietário - formatos de ficheiros com patente ou *copyright*

(fonte: *Wikipédia Consult. Dez. 2007*)

Formato normalizado - Formatos cujas definições e características estão especificadas numa norma, ou é consensualmente utilizado pela comunidade cibernética,

(fonte: Wikipédia Consult. Dez. 2007)

Formato aberto - Formato com especificações públicas para armazenamento, geralmente mantido por organizações de normas não proprietárias e de uso livre de restrições legais.

(fonte: Wikipédia Consult. Dez. 2007)

Gestor - ver Administrador

Horizonte de obsolescência - Prazo contado a partir da data de produção de um objeto digital a partir do qual, não se verificando ações de preservação, a possibilidade deste se tornar obsoleto incrementa numa escala geométrica.

(fonte: Wikipédia Consult. Dez. 2007)

Informação de Arquivo - Informação registada num suporte, e que:

- É produzida e/ou utilizada no âmbito das atividades da organização;
- Serve de prova às atividades dessa organização

Informação de arquivo eletrónica - Conjunto de dados e/ou objetos digitais (p. ex. ficheiros num repositório de um SI) integrados num SI que são produzidos e/ou utilizados no âmbito das atividades da organização (informação de arquivo na aceção verificada na definição anterior) e servem de prova às atividades dessa organização

Integridade - A integridade de um documento de arquivo refere-se a este permanecer completo e inalterado. É necessário que os documentos sejam protegidos contra alterações não autorizadas. As políticas de gestão de documentos de arquivo devem especificar que tipo de adições ou anotações podem ser feitas a um documento depois da sua produção, em que circunstâncias essas alterações podem ser autorizadas, e quem está autorizado a fazê-las. Qualquer alteração autorizada a um documento de arquivo deve ser explicitamente indicada e reconhecível enquanto tal.

(fonte: NP 4438-1, p. 15)

Latência (em inglês *lag*) refere-se ao atrasos que se podem verificar na comunicação entre computadores (internet, por exemplo) e comunicações via satélite. Em computação, *lag* refere-se ao tempo que um pacote de dados leva a completar um percurso de ida e volta entre um computador local e o seu destino.

(fonte: Wikipédia Consult. Dez. 2007)

Metadados - Ver metainformação

Metainformação - (no contexto da gestão de documentos de arquivo) Informação estruturada ou semiestruturada que permite a produção, gestão e utilização de documentos de arquivo ao longo do tempo, assim como nos e através dos domínios em que são produzidos.

(fonte: Wikipédia Consult. Dez. 2007)

Migração - Processo de transferência de um documento de arquivo de um sistema ou suporte de armazenamento para outro, de forma a assegurar a continua acessibilidade à medida que o sistema ou suporte onde se encontrava de degrada ou se torna obsoleto.

Mirror - Procedimento de proteção de dados e de acesso aos mesmos nos equipamentos informáticos. Este sistema implementa-se na tecnologia de RAID 1. Consiste basicamente em ter 2 discos rígidos ligados. Um é o principal e o outro guarda uma cópia exata do principal armazenando qualquer alteração que se faça em tempo real. Desta forma consegue ter-se 2 discos rígidos idênticos e que permite, se tudo estiver bem configurado, que perante a falha do disco principal, o segundo assuma a sua função impedindo a quebra do sistema e a perda de dados armazenados

(fonte: Wikipédia Consult. Dez. 2007)

NFS - (Network File System), é um modelo de sistema de ficheiros, que tem como função centralizar ficheiros num servidor, formando assim uma pasta virtual. É utilizado para sistemas de ficheiros distribuídos numa rede de computadores de área local. Possibilita que diferentes sistemas ligados a uma mesma rede acedam a ficheiros remotos como se fossem ficheiros locais.

(fonte: Wikipédia Consult. Dez. 2007)

Objeto Digital (OD) - Unidade de informação digital que inclui propriedades do objeto e poderá incluir também métodos de execução de operações no objeto.

(fonte Interpares 2, glossary)

Unidade discreta de informação em formato digital.

(fonte RODA Relatório Final)

Obsolescência - a perda de valor que um bem sofre em resultado do progresso técnico ou da evolução.

(fonte: Priberam)

Pacote de aplicações - conjunto de programas de computador que tem por objetivo o desempenho de tarefas de índole prática, em geral ligadas ao processamento de dados, como o trabalho em escritório ou empresarial. A sua natureza é diferente da de outros tipos de software, como sistemas operativos e ferramentas a eles ligadas, jogos e outros softwares lúdicos, entre outros. (Microsoft Office, Open Office)

(fonte: Wikipédia Consult. Dez. 2007)

Perto de Linha (*Near-line*) - Normalmente trata-se de um tipo de armazenamento de dados menos acessível, mas mais barato, do que o armazenamento em linha, sendo útil para o armazenamento de dados de réplicas (cópias de segurança). Um exemplo disso é um conjunto de tapes com tempos de restauro que vão desde segundos até alguns minutos. Normalmente está associado a um aparelho mecânico que move as unidades de suporte do local de armazenamento para o dispositivo de leitura e/ou gravação de dados.

(fonte: *Wikipédia Consult. Dez. 2007*)

Prazo de retenção - Período, previamente estabelecido para a permanência dos documentos junto da atividade produtora, antes do seu destino final.

(*Dicionário de Terminologia Arquivística, adapt.*)

Plano de classificação - Documento que regista o sistema de classes concebido para a organização de um arquivo corrente e as respetivas notas de aplicação, nomeadamente quanto ao âmbito das classes e à ordenação das unidades arquivísticas abrangidas por cada uma, de modo a determinar toda a estrutura de um arquivo.

(fonte: *Manual para a Gestão de Documentos*)

Plano de Preservação Digital - documento estratégico contendo políticas, procedimentos e práticas/atividades para a constituição de uma estrutura técnica e organizacional que permita preservar de forma continuada Objetos Digitais.

Plataforma - 1) Tecnologia de base de um sistema de computador. Dado que os computadores são dispositivos em camadas, compostos por uma camada de hardware a nível dos chips, por uma camada de controladores de hardware e de sistema operativo e por uma camada de aplicação, a camada mais baixa de um computador é normalmente designada por plataforma
2) Na utilização quotidiana, é o tipo de computador ou sistema operativo a ser utilizado

(*Dicionário prático de informática - McGraw Hill*)

Preservação digital - Conjunto de atividades ou processos responsáveis por garantir o acesso continuado e a longo-prazo à informação e restante património cultural existente em formatos digitais.

(fonte: *Introdução à preservação digital*)

Conjunto de atividades desenvolvidas com o fim de aumentar a vida útil dos OD, salvaguardando a utilização operacional, e protegê-los das falhas de suportes, perda física e obsolescência tecnológica;

Conjunto de atividades que promovem a acessibilidade aos conteúdos;

Conjunto de atividades que assistem na preservação do conteúdo intelectual, forma, estilo, aparência e funcionalidade.

Produtor - Entidade que cria os registos, produz e manipula os conteúdos.

Proprietário - Unidade Orgânica responsável pela informação/documento.

RAID - (Redundant Array of Independent Drives ou Disks) - Conjunto Redundante de Discos Independentes, é um meio de se criar uma unidade virtual de armazenamento composta por vários discos individuais, com a finalidade de aumentar a segurança e desempenho através de redundância.

(fonte: Wikipédia Consult. Dez. 2007)

Réplica - ver backup.

Repositório - local centralizado onde se armazena e mantém informação digital, habitualmente bases de dados ou ficheiros informáticos.

(fonte: Wikipédia Consult. Dez. 2007)

Restauro do sistema - Recuperação dos dados do sistema, através das cópias de segurança.

Série documental - Unidade arquivística constituída por um conjunto de documentos simples ou compostos a que, originariamente, foi dada uma ordenação sequencial, de acordo com um sistema de recuperação da informação. Em princípio, os documentos de cada série deverão corresponder ao exercício de uma mesma função ou atividade, dentro de uma mesma área de atuação.

(fonte: Manual para a Gestão de Documentos)

Sistema aberto - Sistema informático concebido com uma Arquitetura tal que permita a junção de componentes de *hardware* e/ou de *software* provenientes de diversos fornecedores.

(Fonte Glossário da sociedade da informação - versão 7)

Sistema de armazenamento - Dispositivo ou conjunto de dispositivos utilizados para armazenar primariamente os dados e a informação produzida.

Sistema de Informação (SI) - sistema informático que, conceptual e fisicamente, se constitui como unidade autónoma ou autonomizável de outros SI. Os SI têm, frequentemente, expressão informacional em bases de dados ou sistemas de bases de dados relacionadas, mas podem também ser constituídos por ficheiros.

Sistema operativo - programa ou um conjunto de programas cuja função é servir de interface entre um computador e o utilizador. (Windows, Linux; MacOS)

Software de base de um computador destinado a controlar a execução de programas, a comunicação entre dispositivos e programas, assegurando as operações de entrada-saída, a atribuição de recursos aos diferentes processos, o acesso às bibliotecas de programas e aos ficheiros, assim como a compatibilidade dos trabalhos.

Nota: O sistema operativo é o *software* mais importante a correr num computador.

(fonte: *Glossário da sociedade da informação - Versão 2007*)

SMB - (Server Message Block) opera como um protocolo de comunicação de rede utilizado.

(fonte: *Wikipédia Consult. Dez. 2007*)

Software livre - segundo a definição criada pela Free Software Foundation, é qualquer programa de computador que pode ser usado, copiado, estudado, modificado e redistribuído sem nenhuma restrição. A liberdade de tais diretrizes é central ao conceito, o qual se opõe ao conceito de software proprietário, mas não ao software que é vendido almejando lucro (software comercial). A maneira usual de distribuição de software livre é anexar a este uma licença de software livre, e tornar o código fonte do programa disponível. O software livre também é conhecido pelo acrónimo FLOSS (do inglês *Free/Libre Open Source Software*).

(fonte: *Wikipédia Consult. Dez. 2007*)

Suporte - Material adequado ao registo de informação. Exemplos de suportes eletrónicos: CD, DVD, tape, Discos rígidos).

(fonte: *Manual para a Gestão de Documentos*)

Tabela de seleção - Instrumento que regista o resultado da avaliação, apresentando-se como uma relação dos documentos de arquivo de um organismo ou administração que fixa os respectivos prazos de conservação em fase ativa e sema ativa e destino final.

(fonte: *Manual para a Gestão de Documentos*)

Unidade binária: - (em inglês Binary unit, ou bit), unidade de medida de transmissão de dados usada na Computação e na Teoria da Informação.

(fonte: *Wikipédia Consult. Dez. 2007*)

Usabilidade - significa que um documento é utilizável ou seja, que pode ser localizado, recuperado, apresentado e interpretado.

(fonte: *NP 4438-1 p. 15*)

USB - Sigla de universal serial bus (bus série universal). Bus série com uma largura de banda de 1.5 Mbps, utilizado para ligar periféricos (modems, impressoras, ratos, teclados, ...) a um microcomputador.

(*Dicionário prático de informática - McGraw Hill*)

Utilizador - ver consumidor

(fonte *OAIS*)

Virtualização de armazenamento (Storage Virtualization) - processo de abstração de armazenamento lógico a partir do armazenamento físico. Este conceito é utilizado atualmente para descrever esta abstração em qualquer camada no conjunto de software e hardware de armazenamento.

(fonte: Wikipédia Consult. Dez. 2007)