

PRESERVAÇÃO COMUM de PATRIMÓNIO DIGITAL



Continuidade digital: relatório final de projeto

Para uma rede comum de preservação de património digital

Apresentam-se as conclusões do trabalho realizado nas etapas previstas do cronograma do projeto. Apontam-se possíveis cenários de continuidade.

(Página intencionalmente deixada em branco)

Ficha técnica MIP:

Título: Continuidade digital: relatório final de projeto

Id: dt_005

Classificação: (MEF) 900.10

Autores: Ana Rodrigues, Francisco Barbedo, Lucília Runa, Mário Sant'Ana

Colaboradores: Ana Franqueira, António Vieira, Cristina Ribeiro, Fernanda Gonçalves, Gabriel David, Luis Corujo, Miguel Azguime, Paula Meireles, Paulo Leitão, Sílvia Ferreira, Teresa Borges, Teresa Gião.

Descritores: 1. Normas. 2. Formatos. 3. Legislação, 4. Terminologia. 5. Representação de informação

Data/Hora: 2015-08-03

Formato de dados: Texto (docx; pdf)

Estatuto de utilização: Restrito ao Grupo de Trabalho

Relação: versão - 0.6



DGLAB, 2015

Conteúdo

0. Sinopse	4
1 Introdução	6
2 Aspetos metodológicos.....	8
3 Quadro regulador	14
3.1 Legislação.....	15
3.2 Terminologias	16
3.3 Normas	21
3.4 Formatos.....	27
4 Valores e práticas estabelecidas	32
4.1 Valores	32
4.1.1 Autenticidade	33
4.2 Práticas	35
4.2.1 Avaliação	35
4.2.2 Outras práticas.....	36
5 Modelo de custos	38
6 Modelo de governação	52
7 Perspetivas futuras	63
8 Anexos.....	69
8.1. Legislação.....	69
8.2 Terminologia	69
8.3. Normas e mapeamentos de normas.....	69
8.4 Formatos.....	69
8.5 Folha de cálculo de custos de serviços de preservação digital	69
Eventos de disseminação organizados pelo GT (SOS digital)	70
Comunicações apresentadas em eventos	70
Documentos de projeto.....	71
Bibliografia consultada	72

0. Sinopse

O presente relatório pode ser lido de várias maneiras:

- Leitura sequencial (provavelmente e justificadamente, a que terá menor adesão!);
- Leitura orientada: no fim de cada secção são apresentadas conclusões sobre o seu conteúdo, as quais constituem um resumo explicativo mas orientado (é a opinião dos autores) sobre a temática tratada;

Nesta sinopse são elencados os pontos que os autores consideraram mais relevantes, sendo indicado o n.º de página onde podem ser encontrados.

Pontos “quentes”

- Se “Recordar é viver” como diz o ditado, o melhor é ir à p. 7 e rever os fundamentos e bases do projeto.
- A consulta dos conceitos aprovados pelo GT do projeto será uma etapa importante para decifrar o restante texto. Podem ser encontrados nas p. 9 e 10.
- Durante o projeto foram elaborados dois inquéritos. Se quiser saber a representatividade dos mesmos vá à p. 13; se quiser saber tudo, mas mesmo tudo, sobre os mesmos, pode consultar os respetivos relatórios em <http://1seminariopreservacaopatrimonioidigital.dglab.gov.pt/projeto-continuidade-digital/documentos-de-projeto/>.
- A estratégia de divulgação dos resultados do projeto assentou em 3 linhas. Quer saber quais? Vá às p. 13 e 14. Se quiser saber tudo sobre as ações de divulgação realizadas, consulte o *Anexo 8.6 Informação adicional*.
- Constatou-se haver por parte das CdP (se não sabe o que é CdP vá ver os conceitos) desconhecimento sobre as leis que enquadram o património digital. Há no entanto alguns aspetos contraditórios. Consulte o enredo na p. 15.
- A questão da perceção das CdP sobre a relevância da autenticidade para a informação digital, apresenta resultados fortemente convergentes. Considerando o possível impacto desta realidade sobre a construção de um repositório tecnológico comum, aconselhamos a sua consulta. Pode ir já para a p. 33.
- Se quiser saber os formatos utilizados por mais que uma CdP vá à p. 28. A variedade de formatos dá uma ideia da pluralidade de informação produzida e do impacto que isso terá num futuro repositório comum (onde todos esses formatos serão ingeridos).
- A avaliação de informação tem semelhanças com a definição do Tempo dado por St.º Agostinho. Toda gente sabe que existe mas ninguém consegue explicar o que é ou como a aplicar... Claro que isto é um exagero destinado a convidá-lo a ir à p. 35, para saber mais.
- Para saber qual a dimensão do universo patrimonial digital declarado vá à p.37. Lá também encontra a estimativa de seu crescimento anual.

- O custo estimado de preservar informação digital depende de muitas variáveis. Estas são analisadas no capítulo 5 e particularmente entre as p. 41 e 47. Descubra ainda porque é que preservar património digital é tão caro...
- Uma das vantagens de uma rede comum para preservar património digital é que quantos mais aderentes houver menos serão os custos que cada um tem de pagar. Para saber como vá à p. 47.
- Se quiser saber quanto lhe vai custar preservar a sua informação digital (será que quer mesmo saber??) use a ferramenta desenvolvida no projeto. Esta pode ser descarregada a partir da p. 69.
- O que acham as diferentes CdP sobre a possibilidade de contribuir financeiramente para a existência de um repositório de preservação comum? Quer saber? Vá à p. 50.
- Gerir uma rede comum não será certamente uma tarefa simples. Se quiser saber os seus direitos e deveres como futuro aderente vá para p. 55.
- Se pretender saber quais os serviços que gostaríamos de proporcionar na rede, salte para a p. 54.
- Na p. 57 encontra um diagrama com um cenário possível de estrutura de rede.
- Para que seja possível uma rede de preservação de património digital partilhada por diferentes CdP, é necessário que existam entre elas pontos de convergência. Que aspetos terão mais peso? Os de convergências, ou os de divergências? Tem curiosidade em saber? Vá diretamente às p. 65 e 66, sobretudo ao diagrama 2.
- A equipa encontrou duas questões problemáticas que exigem mais investigação. Para saber quais são vá às p. 64 e 65.

1 Introdução

1.1 Contexto e relevância do projeto continuidade digital

«É facto facilmente observável a dependência das TIC em praticamente toda a esfera de ação humana. Uma significativa percentagem do que se produz, seja no domínio administrativo, científico ou artístico é mediado tecnologicamente. A fotografia ou o filme são exemplos claros e muito popularizados desta realidade.

Neste contexto, praticamente toda a atividade humana implica a utilização e mediação tecnológicas em menor ou maior grau. A expressão utilizada exaustivamente de “desmaterialização” comporta em si de forma quase inconsciente e reflexa a noção que qualquer mudança a realizar em processos de trabalho implica inevitavelmente a utilização das TIC.

Por este motivo muitos dos objetos que poderão ser considerados como património partilham da particularidade de serem virtuais na medida em que dependem da existência de um sistema intermediário que permita acede-los e utilizá-los. Este sistema é constituído por software e hardware, ambos sujeitos a rápida obsolescência. Do ponto de vista de longevidade, qualidade que normalmente se pretende que as entidades patrimoniais beneficiem, a dependência de um sistema intermediário volátil e descontínuo, tem um impacto dramático.

Sem ações especializadas de preservação tudo pode ser mantido mas nada será acessível e conseqüentemente utilizável devido à obsolescência do sistema intermediário bem como à ausência de ferramentas para o substituir.

A preservação do património digital implica um conjunto de ações concertadas que envolva produtores, detentores e utilizadores da informação, assegurando tarefas diversas e complementares dentro de uma estratégia que apenas pode ter sucesso se for devidamente articulada. Os recursos exigidos para preservar o objeto digital são elevados e exigentes tanto sob o ponto de vista material como de conhecimento especializado. Este fator é tanto mais agravado quanto se verifica falta de recursos no contexto da atual crise económica.

Mas mesmo que nos permitíssemos ignorar esta realidade, poderíamos facilmente encontrar vantagens na criação de estruturas de cooperação que envolvam um conjunto plural de atores e que contribuam num esforço mútuo para assegurar a preservação do património.

A preservação do património digital coloca portanto diversos problemas e dilemas sobre os quais é importante iniciar uma reflexão sistemática.»

Foi com este texto de apresentação¹ que a DGLAB lançou o “1.º Seminário de Preservação Comum de Património Digital” que se realizou a 19 e 20 de setembro 2013.

Na sequência deste evento, e após elaboração do seu “Relatório” e “Conclusões”, foram realizadas várias reuniões - sectoriais e uma geral - com os vários agentes patrimoniais para auscultação sobre proposta de delineamento de estratégia articulada de preservação do património digital. Do interesse e disponibilidade demonstrados, nasceu o projeto Continuidade Digital que, enquanto projeto inovador, pretende também contribuir para a preservação do património digital enquanto memória social e legado para as gerações futuras.

Ao trazer para ordem do dia a questão da preservação do património digital, acredita o grupo de trabalho (GT) obter o reconhecimento da sua relevância e pertinência pelo conhecimento potenciado acerca dela. Ao alinhar a referida temática com os objetivos estratégicos do Secretário de Estado da Cultura (SEC), em matéria de promoção e proteção do património digital, e com a política governamental de racionalização e utilização de processos e métodos decorrentes da adoção de novas tecnologias, espera a obtenção do apoio oficial de entidades públicas, mas também privadas, para a viabilização, concretização e sustentabilidade do projeto em causa.

Objetivo geral do projeto:

- Constituir, na prática, uma rede sustentável de preservação comum de património digital.

Objetivos específicos do projeto:

- Gerir património digital em rede, com base na partilha de recursos e serviços, ou no princípio do trabalho cooperativo em rede;
- Constituir uma rede de atores - pertencentes a diferentes comunidades de prática (CdP)- com responsabilidades definidas e identificadas na preservação do património digital, à escala nacional;
- Recolher critérios para identificação de objetos digitais com estatuto de património;
- Obter contributos diferenciados de cada CdP, mas formalizados e reconhecidos;
- Promover cooperativamente a fruição e acesso ao património preservado.

Passados 18 meses do início do projeto em que se efetuaram 10 reuniões plenárias², ocorreram 11 sessões da iniciativa SOS digital³, foram publicados 5 documentos⁴, pretende-se, através deste relatório, marcar este momento no projeto dando conta do estado atual do mesmo.

¹ O texto está disponível em WWW: <URL: <http://1seminariopreservacaopatrimonioidigital.dglab.gov.pt/wp-content/uploads/sites/19/2014/01/preservacao-comum-de-patrimonio-digital.pdf> >.

² Aceda à equipa de projeto em WWW: <URL: <http://1seminariopreservacaopatrimonioidigital.dglab.gov.pt/projeto-continuidade-digital/equipa-de-projeto/> >.

³ Dimensão de divulgação pública do projeto em WWW: <URL: <http://1seminariopreservacaopatrimonioidigital.dglab.gov.pt/projeto-continuidade-digital/eventos> >.

⁴ WWW: <URL: <http://1seminariopreservacaopatrimonioidigital.dglab.gov.pt/projeto-continuidade-digital/documentos-de-projeto/> >.

Ao longo do presente relatório serão abordados:

No ponto 2 -“Aspetos metodológicos”- o ponto de partida, as várias fases do projeto, a filosofia de investigação, os conceitos e terminologia, o plano e cronograma de trabalho, a forma de obtenção de dados, a estratégia de divulgação de resultados e ainda a forma de desenvolvimento de alguns dos produtos, designadamente o modelo de custos, o modelo de governação e a arquitetura de informação.

Os pontos 3 e 4 são apresentados de forma sucinta, pois foram objeto de desenvolvimento no relatório 1, elaborado no âmbito do projeto e disponível em WWW: <URL: http://1seminariopreservacaopatrimonioidigital.dglab.gov.pt/wp-content/uploads/sites/19/2014/10/Relatorio_Passo1V1.0.pdf >

Os pontos 5 e 6, pelo contrário, encontram-se mais desenvolvidos pelo facto de serem aqui, pela primeira vez, divulgados os resultados da análise efetuada.

2 Aspetos metodológicos

1. Filosofia de investigação

O projeto adotou como base da ação de pesquisa os seguintes princípios orientadores:

1/ metodologia indutiva, ou seja, partir dos dados para a formulação de hipóteses de trabalho adequadas à informação coligida e conhecimento produzido; 2/ “grounded theory”⁵. Trata-se de uma abordagem metodológica que radica na investigação sociológica e que, articuladamente com o método indutivo, propõe o progressivo e dinâmico estabelecimento de caminhos de novos de desenvolvimento à medida que os dados vão sendo recolhidos e analisados. Desta forma não existe um caminho pré definido à partida, mas este vai sendo progressivamente construído à medida que o conhecimento criado aumenta.

Estes princípios orientadores assumem como corolário não existir *a priori* um resultado ou produto esperado do projeto. Ou seja, as possibilidades de cenários e de desenvolvimento futuro vão sendo analisadas e construídas de forma participada, de forma a obter uma base de trabalho aceitável pelas CdP envolvidas.

A razão para esta abordagem teve a ver com a existência pressentida de muitas variáveis de diferente natureza, algumas das quais imprevisíveis e de difícil gestão. Por exemplo uma opção política contrária a opções técnicas eventualmente reconhecidas como válidas.

O projeto foi portanto concebido em duas fases: uma primeira em que, através da recolha, sistematização e análise de dados se pretende obter conhecimento sobre o domínio do

⁵ Glaser, B; Strauss, A- The discovery of Grounded Theory: strategies for qualitative research, ISBN 9780202302607. Ou WWW: <URL: <http://www.groundedtheory.com/>>.

problema e que termina com a formulação de cenários de continuidade possíveis resultantes de acomodações entre os elementos representantes de CdP. Uma segunda fase destinada à execução desses cenários. O presente relatório incide naturalmente sobre os trabalhos decorridos na primeira etapa.

2. Estabelecimento de uma base comum de partida

O primeiro passo foi estabelecer uma base comum de acordo consensual entre os membros do GT. Estes foram convidados pela DGLAB a partir de dois critérios: 1/ a participação no “1.º Seminário para a Preservação Comum de Património Digital”, o qual decorreu no ANTT em Setembro de 2013. 2/ a representatividade de diferentes CdP. Com efeito uma das teses propostas no referido seminário consistia na homogeneidade da informação em ambiente digital no que respeita à sua natureza (codificação binária), diluindo ou mesmo retirando diferenças entre objetos que, tradicionalmente e no mundo analógico, são tratados por CdP de acordo com princípios, métodos e práticas diferenciados.

Considerando a heterogeneidade desejada do GT, era importante estabelecer uma base consensual que formasse um denominador comum a todos os elementos. A partir dessa base ir-se-ia aumentando progressivamente o número de variáveis e requisitos até uma linha em que se verificasse divergência, ou por outras palavras, a não perceção de vantagem adicional. O limite máximo de convergência a que os atores do grupo de projeto chegassem, seria o domínio de implementação, ou seja, o máximo divisor comum aceite por todos os representantes das CdP a partir do qual não se reconhecem vantagens adicionais unanimemente partilhadas.

Esta plataforma foi acordada nos seguintes pontos:

- i. Ser necessário haver vontade efetiva de preservar património digital.
- ii. Preservar património digital significa preservar objetos físicos (ficheiros) que veiculam informação digital.
- iii. Os objetos físicos, independentemente da informação que veiculam, têm atributos comuns.

3. Definição de conceitos

A definição de conceitos e terminologia dentro do GT foi um passo necessário para clarificar e harmonizar o trabalho a desenvolver. A inclusão de termos adicionais, a exclusão de outros inicialmente propostos, bem como a redefinição do significado e entendimento dos mesmos, foram evoluindo em função da recolha de dados e da sua sistematização e análise.

Foram refinados no contexto do GT os seguintes termos:

→ Acesso

- Possibilidade de utilizar e ou reutilizar um recurso patrimonial, geralmente sujeita a regras e condições.

- **Autenticidade**
 - Propriedade de ser autêntico ou seja, possuir os elementos formais definidos pelo ambiente regulador em que o objeto foi produzido e que permitem garantir que este efetivamente é o que pretende ser.
- **Avaliação**
 - Processo de aferição do valor de recursos digitais através dos critérios e métodos adotados por cada CdP e que tem como resultado a seleção para inclusão no programa de preservação digital.
- **Comunidade de interesse**
 - Conjunto de atores que têm algum tipo de interesse sobre o objeto cultural.
- **Comunidade de prática**
 - Conjunto de atores que trabalham, preservam e representam um domínio patrimonial, partilhando normas e práticas comuns.
- **Objetos culturais**
 - Informação produzida e/ou criada enquadrável na definição expressa da Lei de Bases de Património Cultural.
- **Património digital**
 - Todos os recursos de natureza digital, própria ou adquirida, que, sendo testemunhos com valor de civilização ou de cultura, portadores de interesse cultural relevante, devam ser objeto de especial proteção e valorização.
- **Perspetivas**
 - Visão que uma determinada CdP ou utilizador tem sobre o recurso patrimonial. Um mesmo recurso pode ser categorizado de forma diferente, consoante o interesse primordial do observador.
- **Preservação digital**
 - Ações sistemáticas, voluntárias e planeadas realizadas sobre objetos culturais com o propósito de assegurar a sua longevidade, mantendo propriedades básicas de autenticidade.
- **Responsabilidade partilhada**
 - Lógica de responsabilidade comum por todo o património custodiado, independentemente da CdP a que cada grupo de atores pertence.

4. Definição de plataformas e meios de articulação entre os membros do GT

A proposta de organização do GT previa um núcleo executivo de pessoas colaboradoras da DGLAB que teriam funções de coordenação e preparação dos materiais, fossem estes resultantes de recolhas ou de análise. Este núcleo ocupou-se também de toda a logística inerente a reuniões e organização de eventos paralelos. Os membros do GT funcionaram como agentes de debate, trazendo para o grupo as suas diferentes perspetivas relativamente aos tópicos em análise e propostos no plano e cronograma de trabalhos. Serviram ainda como fontes de informação relativamente às suas CdP e tiveram ainda a missão de avaliar e validar os resultados coligidos e analisados.

Relativamente à plataforma de comunicação optou-se pela utilização do Google drive para partilha e edição comum de documentos, bem como plataforma de disponibilização dos documentos de reuniões e atas. Todas as reuniões foram gravadas em áudio, sendo

produzidas atas sintéticas. Foi criado ainda um grupo no Google mail destinado a estimular e promover a troca de mensagens e conversas relativamente a tópicos identificados.

Considerando haver membros do GT que habitam fora de Lisboa, todas as reuniões tiveram uma componente de web conferência assegurada através do *hangout* do Google.

Sempre que solicitada, por pessoas externas ao GT, a adesão ao projeto, foi-lhes facultado acesso ao Google drive e, conseqüentemente, a todos os documentos nele produzidos.

Foi ainda criado um sítio web do projeto⁶ na plataforma *wordpress* da DGLAB, onde foi publicada toda a informação nele produzida.

5. Plano e cronograma de trabalhos

Foi elaborado um cronograma dividido em 7 passos, compreendendo cada um deles um determinado conjunto de atividades a empreender. Este cronograma sofreu alterações ao longo do tempo, que resultaram em junção de atividades em função dos circunstancialismos verificados e da metodologia adotada (ver ponto 1). Foi decidido empreender, sempre que possível, estes passos de forma paralela e não sequencial. O cronograma está disponível no sítio web do projeto, no URL referido no ponto 4. O propósito desta abordagem foi o de encurtar o tempo esperado para terminar a primeira fase do projeto.

Inicialmente foi considerada a subcontratação de empresa externa para proceder ao inquérito do passo 7 e à determinação dos modelos de governação e de custos. No entanto, por razões orçamentais, tal acabou por não ser possível, sendo estas tarefas asseguradas pelo núcleo executivo do GT.

6. Levantamento de dados

Grande parte do trabalho previsto implicou levantamento e recolção de dados, os quais foram posteriormente analisados. Para proceder a este levantamento foram utilizados os seguintes processos:

a) Debate participado com os membros do GT. Neste caso foram solicitados comentários, presencialmente ou remotamente através da aplicação *hangout*, nas reuniões de trabalho. Estes comentários foram captados, conforme atrás referido, em registo áudio, sendo depois escutados e analisados. Procurou-se ainda obter resultados e informação através de mensagens de correio eletrónico enviadas para o grupo. Este meio teve, no entanto, resultados limitados.

b) Levantamento documental e de fontes de autoridade. Recolha não exaustiva, mas o mais representativa possível. Em grande parte foram utilizadas as referências apresentadas pelos membros do GT, as quais serviram de base para a constituição de um corpus de documentos. Em grande parte coube ao núcleo executivo do projeto proceder a identificação de

⁶ WWW: <URL: <<http://1seminariopreservacaopatrimonioidigital.dglab.gov.pt/projeto-continuidade-digital/>>.>

referências na internet, apresentando-as aos membros do GT para posterior validação. Mesmo quando esta não foi concretizada manteve-se a sua inclusão no corpus de análise sempre que tal se considerou justificado. Como critérios de recolha de fontes de autoridade e referências normativas foram considerados:

- I. As referências cedidas pelos membros do GT;
- II. A edição por organismos internacionais de referência de documentos orientadores e normalizadores;
- III. A edição por organismos profissionais de diferentes CdP, não sendo necessariamente organismos normalizadores de orientações técnicas, mas antes coletividades reguladoras de práticas profissionais.

A análise subsequente incluiu análise comparativa dos documentos recolhidos, procurando identificar aspetos, práticas, conceitos e termos comuns às diferentes CdP patentes no corpus de documentos recolhido. Foram também levadas a cabo, com a participação de elementos do GT, experiências empíricas de descrição cruzada, ou seja, aplicação de normas de descrição específicas de uma determinada CdP a material tradicionalmente processado com outras normas.

c) Inquéritos

Foram realizados dois inquéritos: um primeiro fechado e dirigido exclusivamente aos elementos do GT e um segundo dirigido a uma comunidade alargada.

Comunidades de prática	
Inquérito 1	Inquérito 2
Arquivos	Arquivos
	Arquitetura
	Audiovisual
Bibliotecas	Bibliotecas
	Centros de documentação
Cinema	Cinema
Fotografia	
Dados científicos	
Informação clínica	
	Informática
Entretenimento / Multimédia	
Jornalismo	Jornalismo
Museus	Museus
Música	Música
Televisão	
	Publicações periódicas
Rádio	
	Universidades
	Videotecas

Quadro 1 - CdP que responderam aos inquéritos

Inquérito 1	Inquérito 2
Universo = 17	Universo = 800
N.º de respostas = 11	N.º de respostas = 314
% de respostas = 64%	% de respostas = 41%

Quadro 2 - Universo dos respondentes e n.º e % de respostas obtidas

O objetivo desta aproximação consistiu em obter informação, a partir do primeiro inquérito, que permitisse precisar e definir os tópicos e perguntas, assim como a melhor formulação das mesmas, a incluir no segundo inquérito público.

Os resultados do primeiro inquérito serviram ainda como referencial de comparação com os resultados alargados provenientes do segundo inquérito, permitindo assim validá-los, recusá-los ou corrigi-los enquanto tendências estatísticas valorizáveis.

Os resultados dos inquéritos foram tratados com recurso a métodos de estatística descritiva. Não foi utilizada estatística inferencial.

As conclusões gerais destes dois inquéritos foram utilizadas para determinar condicionalismos, classes e operações na configuração e desenho prospetivo de uma rede comum para preservação de património digital (ver ponto 6).

7. Estratégia de divulgação de resultados do projeto

A divulgação do projeto e dos respetivos produtos assentou em 4 linhas:

i) SOS digital. Este evento, inicialmente pensado com um cenário diferente, foi retomado para apoiar especificamente o projeto. As linhas programáticas deste evento de curta duração (normalmente uma tarde) consistiam em: a) publicitar uma determinada componente do projeto e b) convidar especialistas em áreas relacionadas com a preservação de informação digital de forma a obter, sobre esse tópico, esclarecimentos que permitissem aos membros do GT uma reflexão produtiva.

Dentro desta última linha foram organizados três SOS digitais temáticos, um que contou com a presença de Nancy McGovern (Massachusetts Institute of Technology - EUA) que apresentou uma comunicação sobre alguns aspetos específicos de preservação digital. Ainda com Nancy McGovern foi organizada uma reunião de trabalho restrita aos membros do GT e a alguns convidados.

Um outro SOS versou sobre *storage* e um terceiro sobre património fonográfico e sonoro. Foi ainda organizado um SOS especial que contou com alguns especialistas convidados, para além dos membros do grupo executivo do projeto, que pretendeu apresentar e clarificar alguns conceitos básicos de preservação digital. Este último evento decorreu no auditório do ANTT.

ii) Participação em eventos especializados de determinadas áreas científicas e técnicas. Neste domínio foi apresentada uma comunicação no “Encontro Arquivos Científicos”, que decorreu na Faculdade de Ciências Sociais e Humanas da Universidade Nova de Lisboa (organizado pelo Instituto de História Contemporânea da FCSH-UNL e pelo Arquivo de Ciência e Tecnologia da Fundação para a Ciência e a Tecnologia) em 2014; no DLM Fórum, que decorreu em Lisboa em 2014. Foram ainda aprovados uma comunicação e um painel no âmbito do Congresso BAD (Associação Portuguesa de Bibliotecários, Arquivistas e Documentalistas) a decorrer no presente ano de 2015, em Évora.

iii) Publicação dos resultados, relatórios e gravações dos eventos no sítio web do projeto. Relativamente à forma de publicação dos documentos produzidos foi adotada pelo GT a licença cc Attribution-ShareAlike 4.0.

iv) Compilação de recursos bibliográficos de forma a obter uma base de conhecimento teórico representativa do estado da arte no que respeita ao tema e à realidade de cada CdP. Os documentos recolhidos foram colocados no Google drive do projeto para consulta por parte dos membros e convidados do projeto. Presentemente estão também disponíveis no sítio web do projeto.

8. Elaboração de modelos e cenarização

Para o desenvolvimento de alguns dos produtos, designadamente o modelo de custos, modelo de governação e arquitetura de informação, foram utilizadas as aplicações *Cmaps* para a elaboração de modelos conceptuais e o *Modelio*⁷ para elaborar modelos formais através de uma notação mais rigorosa. Foram ainda utilizados, no modelo de custos, ferramentas internacionais de simulação de custos, tendo no entanto sido desenvolvido um protótipo pelo grupo executivo do projeto com recurso a consultas informais externas e à aplicação *Excel*.

O propósito da modelação consistiu na obtenção de diagramas - livres e esquemáticos - que permitissem visualizar cenários admissíveis representando atores e ligações entre os mesmos. Cada modelo foi posteriormente descrito textualmente numa matriz.

3 Quadro regulador

O quadro regulador foi considerado para as seguintes vertentes: legislação, terminologias, normas e formatos utilizados.

Através da análise comparativa dos documentos identificados em cada uma delas, pretendeu-se averiguar da existência de eventuais pontos de convergência e/ou de divergência entre as diferentes CdP.

⁷ Ver Cmaps <http://cmap.ihmc.us/> e Modelio <https://www.modelio.org/>

Há a mencionar que nem sempre foi possível aceder ao conteúdo das terminologias e das normas identificadas, seja por não se encontrarem acessíveis em linha, seja por o acesso às mesmas implicar pagamento.

3.1 Legislação

3.1.1 Procedeu-se, para cada CdP, ao reconhecimento e identificação dos diplomas reguladores da respetiva atividade, tanto a nível nacional como europeu - Regimes jurídicos, Lei de bases, Códigos deontológicos -, assim como de aspetos específicos com especial interesse para a preservação digital, tendo os mesmos sido agrupados em torno dessas duas categorias.

3.1.2 A lista da legislação consultada está disponível no anexo 8.1 do presente relatório.

3.1.3 A informação recolhida através dos inquéritos realizados permitiu detetar um acentuado desconhecimento no que se reporta à regulação do património digital, como pode verificar-se através do gráfico que se segue:

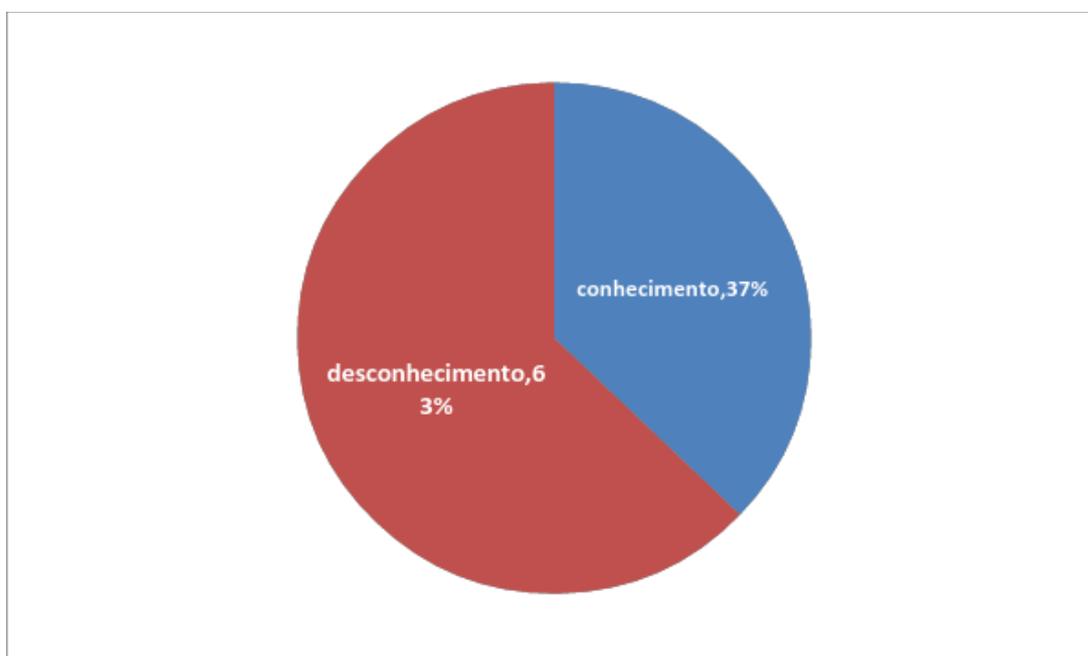


Gráfico 1 - Conhecimento / desconhecimento de legislação que condicione ou determine procedimentos sobre os ODs custodiados

3.1.4 Há no entanto a salientar que estes resultados são de alguma forma contrariados quando, relativamente aos critérios de avaliação de ODs, as respostas ao inquérito demonstram que os diplomas legais, se bem que em conjunto com os regulamentos, são, de entre as oito hipóteses de escolha de fontes de autoridade enunciadas, a mais utilizada por todas as CdP, com exceção da CdP Museus (em que diplomas legais e regulamentos, se bem que em pé de igualdade com outras três opções de resposta, correspondem à terceira fonte mais utilizada) e do conjunto de CdP agrupadas na designação “outras” (em que são, de entre as já referidas oito opções de resposta, a segunda fonte mais utilizada).

3.1.5 Do levantamento efetuado, constatou-se a existência de diplomas:

Com interesse para todas as CdP, a nível nacional e europeu (Lei dos Direitos de Autor e Conexos; Lei do Depósito Legal; Licenças Públicas *Creative Commons sobre dados espaciais*);

- Multidomínio, aplicáveis à atividade jornalística, exercida através da Imprensa, Rádio e Televisão, e Património Cultural;
- Com referências explícitas a preservação de património/digital, designadamente:
 - Art.º 11.º da Lei de Bases do Património Cultural sobre “dever de preservação, defesa e valorização do Património Cultural”;
 - Cap. VII da Lei da Televisão, n.º 1, 2 e 3 do art.º 92.º sobre “conservação do Património Televisivo”;
 - Cap. VII da Lei da Rádio, art.º 83.º sobre “conservação do Património Radiofónico - registos de interesse público”;
 - Portaria 247/2000 - Regulamento de conservação arquivística para os Hospitais e Serviços de Saúde. O processo clínico é de conservação permanente.

3.2 Terminologias

3.2.1 As estruturas vocabulares identificadas constituem o anexo 8.2 do presente relatório.

3.2.2 A fidedignidade da sua produção revelou-se variável: no domínio da Música apenas foi encontrada uma ISBD, ainda em versão *draft*, elaborada na Biblioteca do Congresso (n.º 21 do anexo). Encontram-se, no entanto, inúmeras estruturas terminológicas compulsadas por diferentes autores: pessoas individuais, empresas comerciais (por exemplo a Naxos WWW: <URL: <http://www.naxos.com/>>), blogs, etc.

3.2.3 Algumas das estruturas vocabulares identificadas são muito genéricas - mencione-se, a título exemplificativo, o *EuroVoc: thesaurus multilingue da União Europeia* - enquanto outras são muito específicas - veja-se o caso de *Glossary Filmographic Terms*.

3.2.4 Restrições à sua visualização total determinaram, em parte, a impossibilidade de uma comparação termo a termo. Refira-se, no entanto, que a elevada variabilidade de palavras e termos utilizados (polissemia), bem como a sua diferenciação em duas classes distintas e a existência de formas de apresentação ocultas, que impossibilitaram a comparação, justificam uma metodologia de aproximações semânticas de conceitos e não especificamente de termos.

3.2.5 O ponto de partida da análise foi o vocabulário nuclear estabelecido no documento estratégico do projeto, procurando-se esses termos nas estruturas identificadas. A sua presença, frequência, ou ausência, poderia determinar o peso de que um determinado conceito, independentemente do termo pelo qual é representado, se reveste para uma determinada CdP. O resultado encontra-se sistematizado no quadro que se segue:

	Arquivos	Arte/Arquitetura	Audio-visuais	Teatro	Cinema	Cinema/Som	Fotografia	Inf. Clínica / Científica	Multidomínio	Agricultura	Museus		
Termos base	Dici. Terminolog. Arquivística	AAT Getty	UNESCO Audiovisual	Theatre glossary	Glossary Filmographic Terms	National Film and Sound Archive	CLIP	MESH Medical Subject Headings	UNESCO - Ciências Sociais	União Europeia EuroVoc	União Europeia FAO - AGROVOC	N. Inventário	N.º termos + comuns p/ domínio
Acesso	1	1	1	0	0	1	0	1	1	1	0	0	7
Autenticidade	1	1	1	0	0	0	0	0	0	0	0	0	3
Avaliação	1	1	1	0	0	1	0	0	1	0	0	0	5
Comunidade de Prática	0	0	0	0	0	0	0	0	0	0	0	0	0
Custódia	1	0	1	0	0	0	0	0	0	0	0	0	2
Identificação património digital	0	0	0	0	0	1	0	0	0	0	0	0	1
Infoestrutura preservação digital	0	0	0	0	0	0	0	0	0	0	0	0	0
Modelo de negócio	0	0	0	0	0	0	0	0	0	0	0	0	0
Património digital	0	1	1	0	0	1	0	0	1	0	0	0	4
Preservação digital	1	1	0	0	0	1	1	0	0	1	0	0	5
Certificação e segurança do repositório	0	1	0	0	0	0	0	1	0	0	0	0	2

Direitos de autor	1	1	1	0	1	0	0	1	1	1	0	0	7
Perspetiva de usabilidade	1	1	1	0	0	0	0	0	0	0	0	0	3
Responsabilidade partilhada	0	0	0	0	0	0	0	0	0	0	0	0	0
Arte digital	0	1	0	0	0	0	0	0	1	0	0	0	2
N.º termos comuns p/ domínio	7	9	7	0	1	5	1	3	5	3	0	0	

Quadro 3 - Presença de termos por domínio

Legenda: 1 = presente; 0 = ausente

3.2.6 Relativamente à utilização destas estruturas, e recorrendo à informação recolhida através de inquérito, é possível concluir que existe alguma prática de indexação de ODs, orientada por referenciais internacionais e nacionais, ou ainda por referenciais produzidos internamente por cada entidade, de forma individual, ou disponibilizados na Internet. Em relação a estes últimos (glossários e terminologias), por não ter sido disponibilizado qualquer tipo de identificação e caracterização, não foi possível proceder a análise.

3.2.7 Os instrumentos de indexação mais utilizados são as terminologias e os tesouros, logo seguidas dos glossários, como se pode observar através da análise do gráfico que se segue:

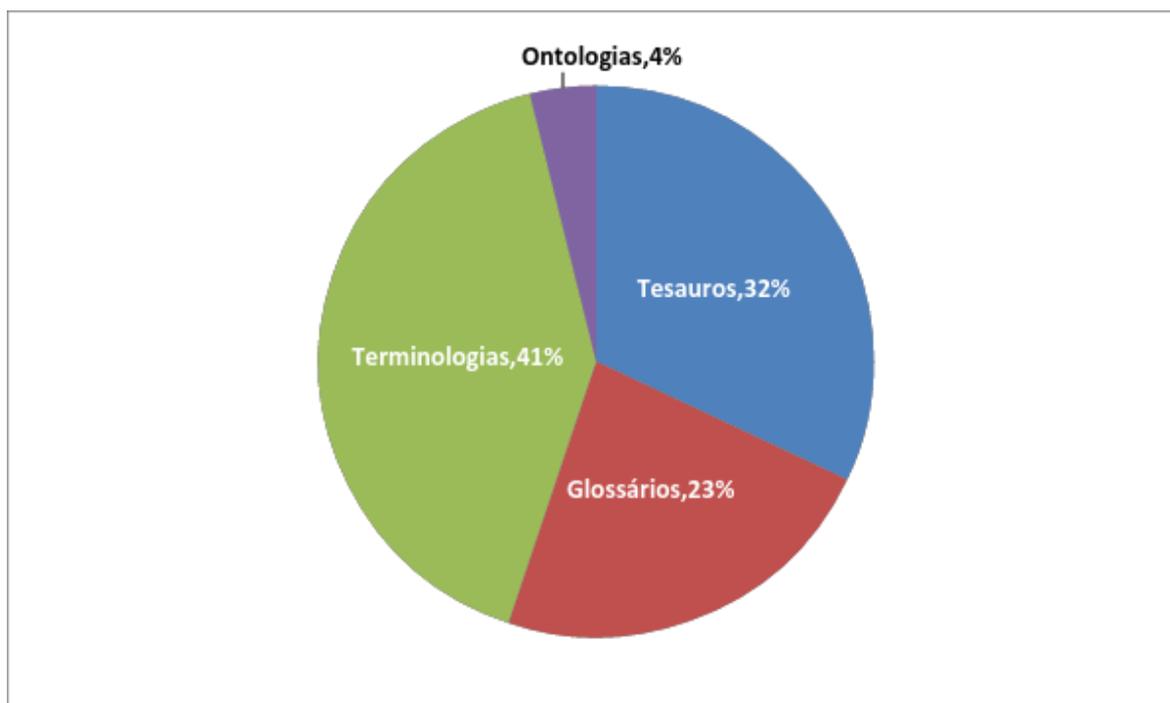


Gráfico 2 - Estruturas vocabulares de indexação utilizadas

3.2.8 Foram identificadas duas grandes classes de estruturas vocabulares:

- a. Desenvolvidas para apoiar diretamente a execução de uma determinada atividade: apresentam termos muito específicos e praticamente exclusivos da CdP a que se destinam. É o caso, por ex., da GEMET Thesaurus (n.º 18 do quadro), relativo a ciências do ambiente. O seu vocabulário é muito “tipado” e orientado para aspetos técnicos ou tecnológicos da atividade a que se referem.
- b. Desenvolvidas para apoiar a descrição ou representação de uma determinada atividade: apresentam uma maior aproximação e similitude de termos. São normalmente produzidas por organizações especializadas na custódia de objetos patrimoniais, para proceder à sua referência e representação descritiva. Nestes casos existe uma semântica subjacente comum, traduzida na existência de alguns termos e conceitos similares.

Aparentemente a diferenciação das formas de representação do mesmo conceito aproxima-se da prática instalada na CdP a que se refere cada estrutura terminológica.

3.2.9 As formas de apresentação destas estruturas são essencialmente duas:

- a. Sistemática, em que o conjunto de termos é apresentado sob a forma de lista alfabética;
- b. Estruturada, assumindo a forma de tesouro, em que são estabelecidas relações hierárquicas ou semânticas entre os termos.

3.2.10 Do ponto de vista da visualização e acesso encontramos diferentes situações:

- a. Pesquisa, sem possibilidade de navegação (*browsing*) ou de descarregar o conjunto dos termos da estrutura;
- b. Visualização direta dos termos, ainda que organizados por grupos alfabéticos;
- c. Apresentação híbrida, com disponibilização das duas possibilidades combinadas.

3.2.11 O número de termos encontrados por domínio encontra-se representado no gráfico que se segue:

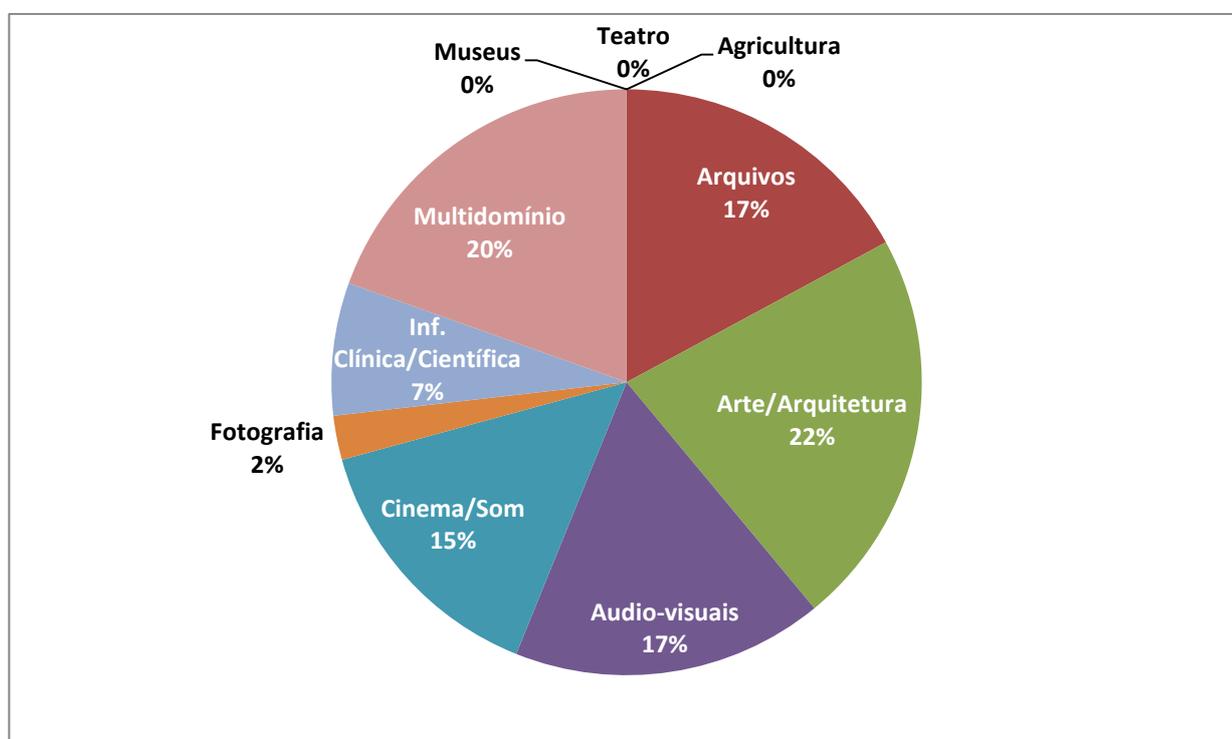


Gráfico 3 - Número de termos por domínio

Destacam-se os termos: *acesso* e *direitos de autor*, seguidos dos termos *avaliação* e *preservação digital*.

3.2.12. A média de número de termos de base utilizados presentes nas estruturas observadas é de 3,7, havendo várias que não apresentam nenhum deles.

O maior número de termos concentra-se nas estruturas vocabulares pertencentes à classe dois, ou seja, orientadas para a representação do objeto, verificando-se a maior concentração em Arte/Arquitetura, Arquivos e Audiovisuais.

Outras estruturas vocacionadas para a execução não contêm quaisquer termos, ou contêm-nos de forma limitada. Sirva de exemplo o vocabulário de termos filmográficos.

3.3 Normas

3.3.1 Foram identificadas normas, regras, guias, orientações, recomendações, etc.:

- Internacionais, emanadas de organismos como: International Council on Archives (ICA), International Federation of Television Archives (IFTA), International Council of Museums (ICM), International Federation of Library Associations and institutions (IFLA), Music Library Association (MLA), Online Audiovisual Catalogers (OLAC), International Organization for Standardization (ISO), EUROPEANA, Archives Portal Europe Network of Excellence (APEX), etc.
- Nacionais, elaboradas por entidades como: Instituto Português da Qualidade (IPQ), Arquivos, bibliotecas e museus nacionais, coordenadores das políticas adotadas nos diferentes setores - Instituto dos Arquivos Nacionais/Torre do Tombo-Direção-Geral de Arquivos-Direção-Geral do Livro, dos Arquivos e das Bibliotecas; Biblioteca Nacional de Portugal; Instituto Português de Museus-Instituto dos Museus e da Conservação-Direção-Geral do Património Cultural; Library of Congress; National Archives of Australia; Association of Canadian Archivists, etc.
- Gerais e específicas.

3.3.2 A lista de normas identificadas corresponde ao anexo 8.3.

3.3.3 A análise comparativa, que incluiu o mapeamento dos elementos constitutivos, foi efetuada apenas para as normas realmente utilizadas pelas diferentes CdP, no sentido de encontrar eventuais correspondências e avaliar possibilidades de mapeamento para uma estrutura comum. Constam do quadro que se segue:

	ISDIAH	EAG	ISAD(G)	EAD2	ISAAR (CPF)	EAC (CPF)	ODA 3	DC	MPEG 7	ISBD	CIDOC	CIDOC- CRM	N In- ven- tário	EBU- TECH 3293
Arquivos	1	1	1	1	1	1	1	1	0	0	0	0	0	0
Audiovisuais	0	0	0	0	0	0	0	0	1	0	0	0	0	0
Bibliotecas	0	0	0	0	0	0	0	0	0	1	0	0	0	0
Inf. clínica	0	0	1	1	1	0	1	0	0	0	0	0	0	0
Muitidomínio	0	0	0	0	0	0	0	1	0	0	0	1	0	0
Museus	0	0	0	0	0	0	0	0	0	0	1	1	1	0
Televisão	0	0	1	0	0	0	0	0	0	0	0	0	0	1

Quadro 4 - Normas selecionadas para análise comparativa

Legenda: 1 = utilizada; 0 = não utilizada

De referir que não foi possível aceder ao texto da norma MPEG 7.

Embora as normas de codificação da descrição utilizadas pelos arquivos tenham sido consideradas como de menor pertinência, apresentam a vantagem de estruturar a informação de forma mais atomizada do que as suas congêneres, o que facilita a abordagem comparativa.

A análise encontra-se documentada de forma mais detalhada no anexo 8.3.

3.3.4 Considerando quer a sobreposição de objetos entre as normas dos diferentes domínios, quer a abordagem de base utilizada em cada uma delas, destaca-se, desde logo, a possibilidade de se considerar três grandes classes: Arquivos, Bibliotecas e Museus -, integrando-se as demais nestas três grandes categorias. Veja-se, a título meramente exemplificativo, o caso do cinema, que utiliza como normas *The FIAF cataloguing rules for film archives* e a *Descrição bibliográfica internacional normalizada (ISBD): edição consolidada*. Distribuem pelas CdP da seguinte forma:

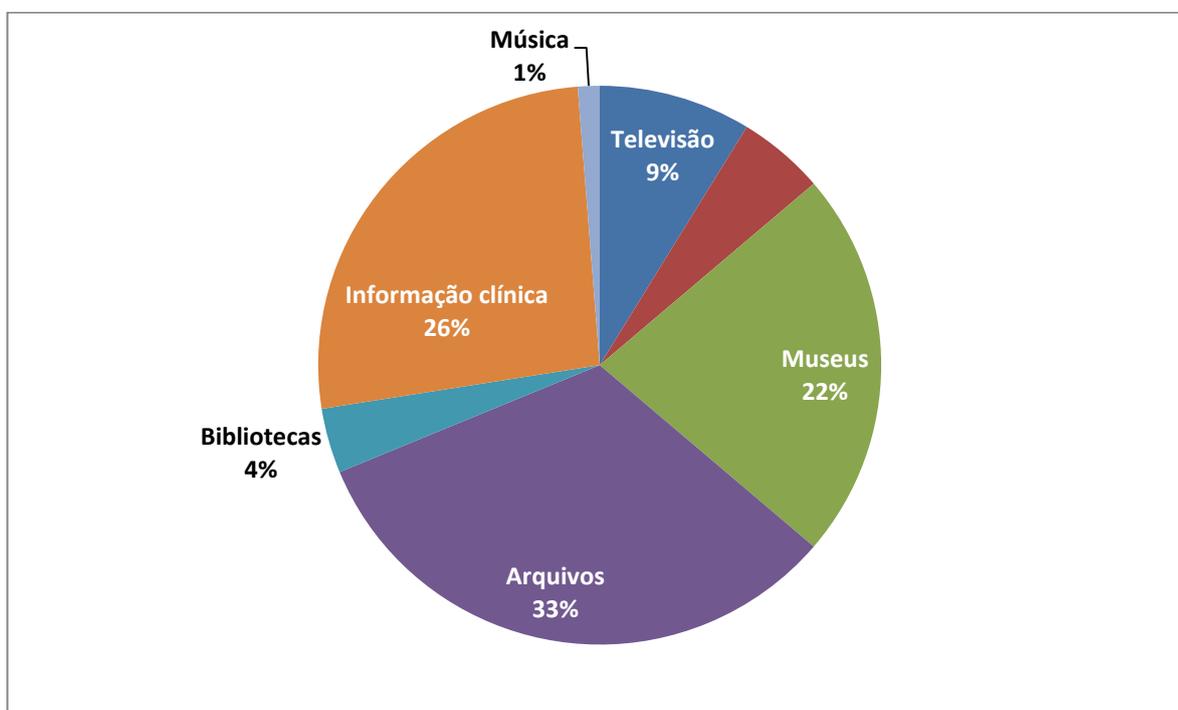


Gráfico 4 - Distribuição das normas por CdP

Do conjunto de normas identificadas quatro são partilhadas por duas CdP com representação no GT, segundo a distribuição esquematizada no quadro abaixo:

Normas partilhadas	CdP
ISAD (G)	Arquivos, Televisão
ISBD	Bibliotecas, Cinema
EDM	Arquivos, Museus
DC	Arquivos, Música
Total	8

Quadro 5 - Normas partilhadas por diferentes CdP

3.3.5 Através de inquérito realizado, destaca-se que uma significativa percentagem de CdP, sobretudo das agrupadas de forma não discriminada, logo seguida dos Arquivos e dos Museus, utiliza práticas não formalizadas para a descrição dos ODs, como aliás se pode verificar através do gráfico que se segue.

Através do mesmo gráfico, percebe-se, no entanto, que para a maioria das CdP - Arquivos, Bibliotecas e Museus - a prática de descrição é normalizada: com recurso sobretudo a normas nacionais e internacionais, e exatamente nesta ordem, no caso das Bibliotecas e dos Museus, e praticamente com o mesmo peso, no caso dos Arquivos.

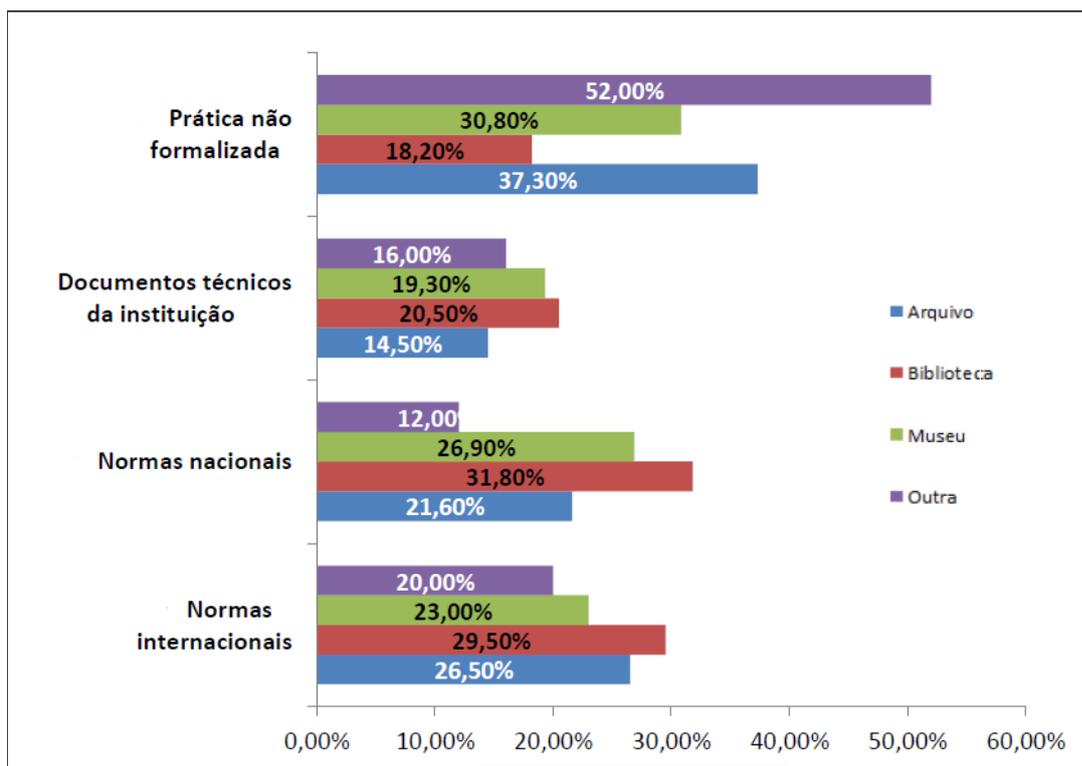


Gráfico 5 - Utilização das normas pelas diferentes CdP

3.3.6 As normas analisadas para Arquivos, Informação clínica, Bibliotecas e Museus, com exceção da CIDOC-CRM, contemplam separadamente a descrição e a codificação da descrição. As normas utilizadas pela Televisão (com exceção da ISAD (G)), pelos Audiovisuais e as transversais (multidomínio), contemplam, na mesma norma, as duas funcionalidades.

3.3.7 A maior parte das normas analisadas centra-se, quanto ao objeto, num domínio patrimonial específico. Constituem exceção a CIDOC-CRM e a Dublin Core, vocacionadas para uma aplicação transversal, multidomínio. Como pode constatar-se através do quadro 4, estão a ser utilizadas pelos Arquivos e pelos Museus, respetivamente.

3.3.8 Constitui exceção o CIDOC-CRM, centrada no evento e orientada para o objeto.

3.3.9 Ainda quanto ao objeto, e socorrendo-nos da terminologia utilizada nas próprias normas, pode ser sistematizado, em cada caso, da seguinte forma:

	Arquivos	Audiovisuais	Biblioteca	Inf. Clínica	Multidomínio	Museus	TV
Objeto	Documentação	Imagens fixas Gráficos Modelos 3D Áudio Vídeo	Recursos publicados	Documentação	Recursos heterogéneos (património cultural)	Objetos e espécies	Recursos: áudio e vídeo
Contexto	Produtores Detentores			Produtores Detentores	Histórico Geográfico Teórico	Histórico Geográfico Teórico	

Quadro 6 - Objeto das normas

3.3.10 Há a salientar, relativamente aos arquivos, o peso assumido pelo contexto de produção, que se materializa na existência de normas específicas para a descrição dos produtores/detentores e das funções. Este aspeto assume também relevância, embora não equivalente, nas normas utilizadas pelos Museus, de entre as quais a CIDOC-CRM se assume como multidomínios. Nas demais normas esta informação é contemplada, mas com muito menor peso e incidência.

3.3.11 Há igualmente a mencionar, para os Arquivos e Informação Clínica, o peso assumido pela descrição multinível, cujas definição e regras são expressamente enunciadas na ISAD (G) e nas ODA I. O modelo hierárquico é igualmente mencionado no domínio dos Audiovisuais. Para as Bibliotecas e os Museus, como pode verificar-se através da análise do quadro 6, prevê-se a descrição dos respetivos objetos como um todo, ou das várias partes que o constituem, dos objetos individuais ou dos conjuntos em que se integram.

	Arquivos	Audiovisuais	Biblioteca	Inf. clínica	Muitidomínio	Museus	Televisão
Objeto	Descrição multinível	Modelo hierárquico: níveis sucessivos	Descrição multinível (parte do recurso/todo o recurso)	Descrição multinível	Classes e subclasses Categorias e subcategorias	Objetos individuais /Conjuntos	

Quadro 7 - Descrição multinível

3.3.12 Quanto à estruturação da informação, Arquivos, Informação clínica, Bibliotecas e Museus, recorrem a zonas ou grupos e, dentro destes, a elementos ou categorias de informação.

Arquivos, Informação clínica e Museus definem a forma como a informação deve ser registada dentro dos elementos ou categorias de informação, indicando regras e convenções específicas para o efeito.

Normas como as ISDIAH, ISAD (G), ISAAR (CPF), ODA, ISBD, CIDOC e Normas de inventário, podem ser classificadas, tendo em conta as suas características, atrás enunciadas, como categoriais.

De entre as normas analisadas existe um outro grupo, composto pelas normas combinatórias como a Dublin Core, a CIDOC-CRM, a EBU-TECH 3293 (ela própria baseada na Dublin Core) e a MPEG 7, constituídas por metadados, esquema de metadados e linguagem de definição de descrição.

3.3.13 O objetivo das normas analisadas é comum, como pode aliás constatar-se através da análise do quadro que se segue:

	Arquivos	Audiovisuais	Biblioteca	Inf. clínica	Muitidomínio	Museus	Televisão
Objetivo	Recuperação, troca e partilha de informação entre sistemas de informação a nível nacional e internacional em cada domínio patrimonial						
	Articulação com outros domínios				Articulação entre domínios patrimoniais: Museus, Bibliotecas, Arquivos. Harmonização de modelos	Articulação com outros domínios	
	Forma autorizada do nome de acordo com outras regras		Aumentar interoperabilidade com outras normas de conteúdo	Forma autorizada do nome de acordo com outras regras		Base de comparação com outras normas nacionais e internac.	
	Relações dos produtores com a doc. arquivo e com outros recursos			Relações dos produtores com a doc. de arquivo e com outros recursos			

Quadro 8 - Objetivo das normas

Para a maior parte delas a preocupação centra-se fundamentalmente na recuperação, troca e partilha de informação entre diferentes sistemas, centrados num domínio patrimonial específico. Para as multidomínio, essa preocupação é assumida de forma transversal. Todas elas, no entanto, assumem a necessidade de articulação com outros domínios: no caso dos Arquivos, essa preocupação traduz-se na total autonomização da descrição dos produtores /

detentores, que para além de facilitar a partilha da informação, permite o estabelecimento de relações com a descrição da documentação por eles produzida / detida, ainda que dispersa por diferentes entidades detentoras, mas também com outros tipos de recursos, nomeadamente biblioteconómicos, museológicos, etc.

No caso das Bibliotecas, a ISBD assume, de forma expressa, a preocupação de aumentar a interoperabilidade com outras normas de conteúdo. No caso dos Museus, na CIDOC e nas Normas de inventário, é assumida a preocupação de fornecer uma base de comparação com outras normas nacionais e internacionais. Por sua vez a CIDOC-CRM assume uma vocação de harmonização de diferentes modelos de descrição.

3.3.14 Centrando-nos na análise dos elementos de informação / classes e das respetivas correspondências entre as diferentes normas, com base no mapeamento que corresponde ao anexo 8.3, pode concluir-se o seguinte:

- As normas analisadas, e centrando-nos primeiro nas categoriais, pese embora a especificidade que decorre dos diferentes objetos para os quais foram elaboradas, são bastante equivalentes.
- Verifica-se que os elementos ou as categorias de informação podem estar preparados para o registo mais ou menos atomizado dos diferentes tipos de informação. Este aspeto - maior ou menor número de elementos para o registo da informação - nem sempre parece estar relacionado com o grau de relevância atribuído à informação a registar. Veja-se, a título exemplificativo, o caso do *Título*: as normas arquivísticas preveem a obrigatoriedade do registo de um *Título*, pelo que disponibilizam um único elemento de informação, sendo os restantes, caso existam, remetidos para o elemento de informação *Notas*. Bibliotecas, Museus e Televisão preveem diferentes elementos para registar esta informação. De igual forma está prevista, nos Arquivos, a indicação do *tipo de Título* registado, sem que exista um elemento de informação específico para esse fim, ao contrário do que se verifica para os Museus, que preveem, para além do elemento *Title*, o elemento *Title type*.
- Algumas das especificidades detetadas prendem-se com o âmbito das normas. A título meramente exemplificativo, veja-se o caso das Bibliotecas, que contemplam elementos de informação específicos para *Dados matemáticos (recursos cartográficos)* e *Apresentação musical (música notada)*. Este tipo de informação é igualmente relevante no domínio dos Arquivos. O que se verifica é que tanto as normas internacionais do ICA como as normas portuguesas sobre elas desenvolvidas são gerais, não contemplando documentos especiais. Em caso de necessidade este tipo de informação é registado em elementos de carácter mais genérico.
- Existem outras especificidades nas normas analisadas que não decorrem do respetivo objeto. Sirvam de exemplo os elementos destinados ao registo da informação relativa à atual localização dos objetos e espécimes, nos Museus, que não tem paralelo nas demais normas analisadas. Embora seja igualmente considerada relevante no domínio dos Arquivos, é objeto de registo

autonomizado. Verifica-se exatamente o mesmo tipo de situação para as intervenções de conservação e restauro.

- Por outro lado, há a mencionar a existência de elementos de descrição vocacionados para informação equivalente mas que, pelas diferentes características dos objetos a que se reportam, assumem especificidade. Sirvam de exemplo *Fonte imediata de aquisição ou transferência* e *Acquisition method*, respetivamente no domínio dos Arquivos e dos museus, já anteriormente mencionados, por contraposição a *Modalidades de aquisição*, nas bibliotecas, que se destina a acomodar informação relativa ao preço ou a outras modalidades de aquisição, se o recurso não estiver para venda.
- Existem, no entanto, elementos especialmente vocacionados para determinados tipos de objetos. É o caso, nas Bibliotecas, dos que constituem a *Zona de Edição* ou a *Zona de publicação, produção, distribuição, etc.*, esta última com correspondência também nas normas utilizadas pela Televisão.
- Quanto às normas combinatórias, a utilização da Dublin Core encontra-se testada no âmbito de vários portais nacionais e internacionais. Quanto à CIDOC-CRM, para além do mapeamento disponibilizado no anexo que corresponde ao anexo 8.3, foram realizados testes práticos, no sentido de determinar da sua maior ou menor adequação à representação dos objetos das diferentes CdP consideradas, tendo-se concluído que, embora subsistindo dúvidas de aplicação prática, o resultado dos testes foi positivo⁸.

3.3.15 Relativamente à Informação clínica, foram identificadas várias normas específicas, tendo sido incluída, no anexo que corresponde ao separador *normas*, a referenciada como de utilização prática: OpenEHR Foundation - *Open European Health Record*. No sentido de perceber até que ponto, e dada a especificidade apontada para a informação clínica e, mais precisamente, para os processos clínicos, existe ou não correspondência entre ela e as outras normas identificadas, e para apurar a possibilidade de mapeamento para uma estrutura comum, foram analisados casos práticos - processos clínicos concretos, devidamente enquadrados no respetivo contexto de produção. Da análise efetuada foi possível concluir que, apesar da sua complexidade, as ODA respondem às necessidades de descrição, quer no que respeita aos níveis de descrição necessários, quer no que se reporta aos elementos de informação para acomodar a informação necessária.

3.4 Formatos

Os formatos que serviram de base ao presente estudo foram os referidos no sítio web The Digital Formats Web site da Biblioteca do Congresso norte-americano⁹, acrescentados dos contributos ad hoc de colaboradores do projeto. Estes formatos, independentemente do fim

⁸ *CIDOC Conceptual Reference Model (CRM): análise geral e análise comparativa com outros esquemas de metadados*. Lisboa: DGLAB, 2014. Documento interno.

⁹ <http://www.digitalpreserv47+47+ation.gov/formats/fdd/descriptions.shtml>

para que foram criados (acesso ou preservação), foram ainda enriquecidos com as referências que as CdP entenderam necessárias para espelhar a sua realidade.

3.4.1 Número total de formatos usados:

Foi referido um total de 68 formatos utilizados pelas diferentes CdP:

AAF	DNG	HTML	MPEG-4 Video Encodings	RealMedia
AC-3	DOC	ID3	MSG	RIFF
AI	DOCX	Indeo	MXD	SQL
ASF	DPX	ISO_8211	MXF	TIFF
AUTOCAD	DTS	ISO_image	MySQL	Uncompressed video encodings
AVI	DV	JPEG	Oracle	WAV / WAVE
BigTIFF	DWF	JPEG 2000 Encodings	PCM	WM (Windows Media)
BWF	DWG	JPEG 2000 File Formats	PDF	XHTM
Cinepak	ESRI Shapefile / SHP	MDB	PNG	XLS
CSV	FLAC	MP2	Postgres	XLS (linux)
DGW	Flash SWF, FLA, FLV	MP3	PPT	XML
DICOM	H.26n ITU-T video encoding standards	MPEG-1	PSD	ZIP
Digital Betacam, AKA Digibeta or D-Beta	HDCAM	MPEG-2	QuickTime	
DivX	HTM	MPEG-4 File Formats with Encoded Bitstreams	RAW	

Quadro 9 - Formatos utilizados pelas CdP

Número total de formatos por CdP:

A CdP Arquivos utiliza 57 diferentes formatos; a CdP Bibliotecas 53 formatos; a CdP Museus 46 formatos; e a CdP “Outros” refere 50 formatos.

Categorias de formatos referidos em absoluto:

A totalidade das CdP utiliza formatos de todas as categorias previstas pelo sítio web da Biblioteca do Congresso norte-americano¹⁰:

- Imagens fixas;
- Som;
- Texto;
- Genérico;
- Imagens em movimento;
- Datasets;
- Arquivo WEB;
- Geoespacial.

Formatos partilhados por mais do que uma CdP:

Apurou-se um total de 54 formatos utilizados por mais do que uma CdP, dos quais 16 se distribuem pelas CdP elencadas:

	HTML	RAW	AC-3	ISO_8211	MXD	AI	MXF	DWG	BigTIFF	ESRI Shapefile /SHP	MDB	PCM	BWF	H.26n ITU-T	DNG	HTM
Arquivos		x	x	x	x		x	x	x	x	x	x	x	x	x	
Bibliotecas	x	x		x	x	x		x		x	x				x	x
Museus	x					x			x							x
Outros		x	x		x		x	x		x	x	x	x	x		

Quadro 10 - Formatos partilhados pelas CdP

¹⁰ Anteriormente referido: WWW: <URL: <http://www.digitalpreserv47+47+ation.gov/formats/fdd/descriptions.shtml> >

Formatos usados por todas as CdP:

Verificou-se a existência de 38 formatos utilizados por todas as CdP:

AAF	FLAC	MPEG-4 File Formats with Encoded Bitstreams	SQL
ASF	Flash (SWF, FLA, FLV)	MPEG-4 Video Encodings	TIFF
AVI	HDCAM	MySQL	Uncompressed video encodings
CSV	JPEG	Oracle	WAV/ WAVE
Digital Betacam, AKA Digibeta or D-Beta	JPEG 2000 Encodings	PDF	WM (Windows Media)
DivX	JPEG 2000 File Formats	PNG	XLS
DOC	MP2	Postgres	XLS (linux)
DOCX	MP3	PSD	XML
DV	MPEG-1	QuickTime	
DWF	MPEG-2	RealMedia	

Quadro 11 - Formatos utilizados por todas as CdP

Conclusões

A análise comparativa efetuada nas diferentes vertentes do quadro regulador, aponta para um conjunto de pontos comuns que importa destacar.

Legislação:

Da legislação analisada a *Lei de Bases do Património Cultural* define requisitos de preservação aplicáveis transversalmente a informação de valor patrimonial, sendo definido como tal uma significativa diversidade de informação inserível em diversas domínios culturais e produzida por diferentes CdP. Legislação específica, como a *Lei da Televisão e da Rádio*, aponta igualmente para a necessidade de preservação, embora não seja especificamente referido o suporte digital. A questão dos direitos de autor é extensível a

todos os domínios culturais, sendo uma questão estruturante que surge com particular relevância quando aplicada em ambiente digital.

Terminologias:

As terminologias analisadas apontam para uma aproximação sempre que o foco das mesmas é a representação do objeto. Divergem, no entanto, quando estas estruturas se centram no aspeto executivo e de apoio à atividade desenvolvida. Esta constatação parece lógica e não deverá ser particularmente valorizada, uma vez que a preservação de património digital incide claramente sobre o primeiro foco, considerando que, em princípio, os objetos preservados não se encontram já operacionalmente ativos.

Dentro das terminologias orientadas para a representação do objeto há conceitos partilhados, embora possam surgir de forma lexicalmente diferente.

Normas:

Normativamente é de salientar a existência de elementos de informação básicos comuns, independentemente do domínio cultural em que são aplicados. Alguns testes efetuados permitiram realizar descrições ou catalogações “cruzadas” utilizando uma norma aplicável a objetos provenientes de diferentes domínios culturais. A necessidade de contextualizar o objeto patrimonial, bem como a representação hierárquica, tradicionalmente consideradas como singularidades arquivísticas, manifestaram-se presentes noutras CdP como os Museus e as Bibliotecas. As normas combinatórias revelam-se suficientemente poderosas para descrever qualquer objeto, independentemente da via de análise adotada (centrada no evento ou no objeto). Na área do Jornalismo e da Música as observações realizadas permitem concluir haver características passíveis de ser representadas com recurso às normas analisadas. Podemos assim concluir ser possível interoperabilidade entre os vários esquemas descritivos utilizados sem perda de sintaxe ou semântica. A utilização de normas diferentes de acordo com as características do objeto a representar, bem como as necessidades específicas do público, não é impeditiva, portanto, do mapeamento de elementos semanticamente idênticos para uma estrutura única.

Formatos:

Os formatos referidos pelos membros do GT não apresentam características de especial singularidade, sendo utilizados correntemente. Os que apresentam maior especificidade são os utilizados na Comunidade médica (DICOM) e na Televisão, com formatos de tapes digitais normalmente utilizados em contextos de captação profissional de imagens em movimento (Digital Betacam, AKA Digibeta or D-Beta, HDCAM). Destaca-se a quase unanimidade da utilização da metalinguagem XML, aspeto relevante se considerarmos que esta metalinguagem fundamenta muitos esquemas normalizados empregues na preservação de ODs.

Foram referidos poucos formatos na categoria de “datasets”, situação que provavelmente se alterará com a presença da CdP Dados científicos e ainda com a exploração posterior dos formatos utilizados pela Informação clínica, onde a informação clínica do doente, nas

diversas vertentes administrativa, técnica e assistencial é normalmente armazenada em bases de dados. A heterogeneidade identificada dos formatos utilizados permite confirmar a necessidade de um controlo sólido dos mesmos, bem como da respetiva evolução de mercado, e ainda de identificação de formatos aceitáveis do ponto de vista de utilização e preservação continuada.

Nalguns casos, indicados no texto, os resultados obtidos pelo segundo inquérito mais alargado vieram corrigir as observações inicialmente avançadas no primeiro inquérito.

4 Valores e práticas estabelecidas

Foram identificados valores partilhados por diferentes CdP. Por valores entendemos um conjunto de princípios que presidem, ou que pelo menos informam de maneira relevante, as práticas das diferentes comunidades. Estas últimas foram igualmente objeto de averiguação no sentido de procurar determinar os procedimentos mais regularmente empregues.

4.1 Valores

De entre os princípios subjacentes à informação (digital) consideramos como mais justificativo de obter dados:

- a perceção de cada CdP sobre o grau de relevância da autenticidade da informação digital.

Conforme referido na secção sobre metodologia foram elaborados dois inquéritos:

- um dirigido exclusivamente aos elementos do GT, visando - por se tratar de uma primeira abordagem - recolher informação para definir tópicos e reformular perguntas a incluir em inquérito posterior (inquérito 2 do passo 2);
- outro, mais abrangente, de carácter público, destinado a uma comunidade mais alargada de respondentes pertencentes a várias CdP, atuando em entidades do sector cultural e criativo, suscetível de validar ou rejeitar as tendências expressas no inquérito anterior.

Os inquéritos referidos¹¹ pressupuseram dois tipos de abordagem de análise descritiva:

- Global dos dados fornecidos em cada resposta;
- Cruzada dos dados de cada CdP com cada uma das respostas.

No âmbito do inquérito 2 do passo 2, foi quantificado, com fundamento em métricas comparativas, fornecidas por cada CdP - inquérito 1 do passo 2, em maio de 2014 - uma base de análise das variáveis propostas para cada aspeto - Autenticidade e Avaliação - tendo em

¹¹ WWW: <URL: <http://1seminariopreservacaopatrimonioidigital.dglab.gov.pt/projeto-continuidade-digital/documentos-de-projeto/> >.

vista determinar graus de relevância - das propriedades e dos elementos externos de contexto de ODs e de aspetos relacionados com avaliação, designadamente indicação precisa de critérios de inclusão e de exclusão de universo patrimonial e respetiva fonte de autoridade

No inquérito alargado - passo 6 -, as variáveis referentes à Autenticidade coincidentes com as questões do inquérito anterior (Q2-Q9), foram ampliadas passando a abordar situações adicionais como modificações dos ODs e admissibilidade das mesmas. As variáveis relativas a Avaliação mantiveram-se, tendo no entanto sido excluída a lista de critérios possíveis de avaliação, por se ter considerado um fator suscetível de influenciar as respostas.

Neste domínio, os resultados obtidos pelos dois inquéritos efetuados foram tendencialmente concordantes, embora se tenham verificado alguns desvios.

As principais conclusões são a seguir apresentadas.

4.1.1 Autenticidade

A perceção de autenticidade foi analisada sob duas perspetivas: 1/a da sua relevância; 2/a da admissibilidade de modificações e respetivo tipo de aceitação por motivo de ações de preservação dos ODs custodiados sem perda de autenticidade.

Relativamente à primeira, - Relevância -, nos dois inquéritos, a abordagem visou analisar os seguintes aspetos da questão:

- a) cada uma das propriedades que contribuem para a autenticidade da informação contida no OD - Identidade, Identificação, Integridade, Conteúdo, Contexto, Usabilidade e Estrutura - e que lhe garantem a capacidade de ser aquilo que pretende ser, ou seja, autêntico;
- b) cada um dos elementos externos - Tecnológico, Procedente, Documental, Jurídico, Administrativo e Procedimental - que integram o contexto de produção da informação digital e que contribuem igualmente para a sua autenticidade.

A perceção global e por CdP - resumida em quadros e representada graficamente nos respetivos relatórios estatísticos dos passos 2 e 6 -, permitiu concluir relativamente ao aspeto referido na alínea a), que:

- Existe uma forte perceção de relevância no conjunto das variáveis em análise: propriedades e elementos externos do contexto de produção;
- As propriedades com resultados mais prevaletentes, mas não significativos - ainda que com escalas de avaliação e valores diferenciados -, foram o Conteúdo e a Identidade;
- A propriedade com resultados menos prevaletentes foi a Estrutura;
- Os elementos externos mais valorados, embora não de forma significativa, foram os seguintes: Tecnológico, Documental e Procedente; o menos pontuado foi o Administrativo;

- Existe uma clara tendência para reconhecer a importância da Autenticidade dos ODs custodiados.

Relativamente ao aspeto da admissibilidade de modificações e tipo de aceitação das mesmas foram equacionados dois aspetos: a/ a aceitação ou não de modificações a ODs e b/ quais as modificações toleradas.

A análise da Autenticidade sob a perspetiva da auscultação da aceitação de modificações dos ODs, provocadas por ações de preservação, consistiu, primeiramente, em testar quatro variáveis do problema:

- a) São admissíveis quaisquer alterações, desde que documentadas;
- b) São admissíveis quaisquer alterações, ainda que não documentadas;
- c) São admissíveis algumas alterações, considerando casos específicos, e desde que documentadas;
- d) Nenhuma alteração é admissível.

Os resultados globais referentes a estas variáveis - conforme representação gráfica dos mesmos no relatório estatístico do passo 6 - permitiu perceber qual o entendimento dos respondentes acerca das referidas modificações, constatando-se, a esse respeito, uma clara convergência de opinião pela aceitação da situação expressa em c).

Os dados permitiram ainda concluir que:

- A maioria dos respondentes acredita não ser possível preservar ODs sem admitir alterações;
- Existe uma quase total discordância relativamente à aceitação de quaisquer alterações não documentadas;
- As alterações deverão ser alvo de processo documentado, pressupondo como tal, controlo.

As conclusões a retirar sobre o entendimento das CdP poderão ser sintetizadas da seguinte forma:

- A tendência revela que todas as CdP admitem apenas algumas alterações de casos específicos e desde que documentadas;
- Com exceção da CdP de Biblioteca, todas as restantes CdP entendem que as alterações deverão ser criteriosas, não bastando, por conseguinte, a mera documentação das alterações efetuadas no âmbito das ações de preservação;
- Todas as CdP consideram que as alterações deverão ser alvo de processo documentado, e como tal, controlado;
- Todas as CdP acreditam não ser possível preservar ODs sem aceitar alterações.

Relativamente aos tipos de modificação tolerados foram propostas modificações a 3 camadas diferentes do OD: Conteúdo, Estrutura e Contexto.

Constatou-se que a variável mais consensual e definidora da tendência foi a alteração ao Conteúdo revelando significativa discordância no que se refere à aceitação de alterações.

Os resultados parciais permitiram concluir que:

- Há alterações que são **unanimemente rejeitadas** por todas as CdP, como as do Conteúdo;
- Há alterações que são **unanimemente aceites** por todas as CdP, como as da Estrutura;
- Há alterações que são **maioritariamente aceites** pelas CdP, como a da informação de Contexto.

4.2 Práticas

Relativamente a práticas procurou-se obter informação sobre os procedimentos empregues por diferentes comunidades no que respeita a aspetos específicos de preservação de informação digital. Para efeitos de organização deste tópico foram criadas duas classes:

- A avaliação da informação digital, nomeadamente se existe uma prática sistemática da mesma e se esta é baseada em critérios próprios ou universais adotados pelas diferentes CdP.
- Outras práticas:
 - Programas de preservação, nomeadamente se estes existem de forma oficial e planificada nas instituições;
 - Sistemas de armazenamento e nomeadamente se estes são adaptados à função de preservar informação digital;
 - Utilização de sistemas TIC específicos para a gestão e preservação de ODs;
 - Custódia de ODs, nomeadamente em que medida as CdP reconhecem custodiar património digital.

4.2.1 Avaliação

Foram consideradas as questões sobre a utilização da prática de Avaliação e quais as fontes que os respondentes utilizam para utilizar e aplicar critérios para a sua realização.

Em termos globais, as respostas permitiram concluir que:

- A maioria dos respondentes “Utiliza sempre critérios de avaliação”, porque tem essa prática instituída no seu ambiente profissional;
- A maioria dos respondentes manifestou discordância relativamente ao facto de nunca se fazer avaliação e é de opinião de que esta é uma prática necessária;
- Todas as CdP confirmam as duas tendências apuradas globalmente;
- A CdP que mais utiliza critérios de avaliação é a de Arquivos, enquanto que Bibliotecas e os Museus são os que menos os utilizam.

A análise dos dados globais dos dois inquéritos permitiu concluir:

- Grosso modo, os critérios de avaliação dos ODs baseiam-se maioritariamente em referenciais internacionais e nacionais;
- É positivo constatar a existência de utilização de fontes de autoridade de critérios de avaliação pela ajuda e segurança que podem conferir a eventuais previsões e estimativas de crescimento anual de património digital a preservar;
- Existe uma tendência para utilizar leis e regulamentos;
- As normas internacionais e nacionais, bem como os documentos técnicos nacionais, também são significativamente utilizados;
- Os documentos técnicos internacionais são menos utilizados;
- A inexistência de critérios e de práticas não formalizadas é ainda uma realidade, indiciando falta de políticas de gestão de documentos e também de gestão de património digital.

Perceção por CdP

Os dados parciais confirmam a tendência dos resultados globais:

- As leis e regulamentos são a fonte de autoridade mais utilizada por todas as CdP;
- As CdP que utilizam mais fontes de autoridade são a de Arquivos e Bibliotecas - sobretudo normas internacionais e nacionais;
- A CdP que mais utiliza prática não formalizada ou não usa qualquer critério é CdP Outra.

4.2.2 Outras práticas

- i. Existência de programas de preservação formalizados. A maioria dos respondentes admitiu não ter programas de preservação formalizados nas suas instituições, registando-se uma tendência que recai em programas informais, i.e., sem existência documentada e formal, e ainda em ações de preservação esporádicas. Destas, a mais frequente é o backup de informação. De uma maneira fortemente consolidada todas as CdP mantêm algum tipo de ação de preservação sobre informação digital.
- ii. Sistemas de armazenamento utilizados. As respostas prevaletes remetem para a utilização de servidores e sistemas de ficheiros e dispositivos externos como DVD para armazenar informação digital. Admitindo a representatividade desta situação é possível identificar um claro obstáculo à qualidade e segurança da preservação de informação digital empreendida atualmente. O armazenamento é, de resto, um dos maiores pontos críticos, facto ao qual não deve ser alheio o elevado custo de sistemas dedicados.
- iii. Utilização de plataformas dedicadas para gerir objetos digitais. Neste domínio a reposta dominante foi a instituição possuir um sistema, embora não dedicado à

preservação e gestão de informação digital. No entanto, apesar de ter sido fornecida explicação sobre o que se entendia por sistema dedicado, é possível ter havido confusão com sistema de gestão de documentos ou com sistema dedicado a produção e disseminação de conteúdos.

- iv. Custódia de objetos digitais. A maioria de respostas refere existir custódia de objetos considerados como patrimoniais. Há uma percentagem com algum significado que refere o contrário, a qual corresponde sensivelmente a idêntica percentagem de respondentes que referiram não realizar qualquer tipo de avaliação. Nestas circunstâncias a declaração de não custódia de objetos digitais com valor patrimonial pode advir do facto de não ter sido feita avaliação. Ainda neste tópico constatou-se que estes objetos digitais sob custódia resultam, em partes idênticas, de produção original digital e de processos de transferência de suporte.

Procurou-se ainda obter uma perspetiva global quanto à dimensão do universo custodiado e ao seu crescimento prospetivo. A quantificação das existências de informação digital são, de forma algo incompreensível, métricas que praticamente as instituições não mantêm, ou mantêm apenas de forma parcelar. Os comentários a esta questão são um exemplo interessante dessa situação. Assim, as respostas a esta questão deverão ser consideradas com prudência. Os valores questionados eram totais, não sendo solicitada a discriminação por tipos de formatos. Os valores reportados em diferentes unidades binárias foram somados e convertidos na unidade binária superior - por exemplo: havendo 1200 MB, estes foram convertidos em 1,17 GB (1200/1024). Algumas respostas foram fornecidas em valores absolutos, sendo referido o n.º de objetos; por ex., 12000 imagens digitais. Embora estes dados sejam apresentados no relatório estatístico, são omitidos no presente documento.

No inquérito estatístico foram reportados c. 2,35 PB de informação digital custodiada. A estimativa de crescimento aponta para c. de 20 TB por ano de nova informação adicionada. Aparentemente verifica-se uma discrepância entre o volume total de informação declarada e estimativa de crescimento, bastante mais modesta. Acreditamos que o elevado valor da primeira variável se deve a respostas de produtores que por si só detêm grande quantidade de informação, como o caso dos Arquivos de Televisão. Nestas circunstâncias este valor deve ser considerado como um *outlier*, pois um só repositório inflaciona significativamente o resultado final.

5 Modelo de custos

A definição de um modelo de custos para a preservação de ODs patrimoniais no âmbito de uma rede comum tem como objetivo não apenas calcular, em abstrato, os custos envolvidos no conjunto de atividades de preservação digital mas, e antes de mais, apurar os custos de preservação digital num conjunto de situações concretas, identificadas e caracterizadas, em diferentes tipos de cenários:

- a) Quando empreendida a título individual, enquanto solução desenvolvida pelas próprias entidades ou contratualizada a empresas especializadas na área¹²;
- b) Quando empreendida no contexto de uma rede comum de preservação, com base numa plataforma tecnológica comum com custos partilhados¹³.

A comparação entre estes dois cenários constitui o modelo de negócio que permite avaliar o custo/benefício e o custo de oportunidade inerentes a cada um dos cenários.

Um dos aspetos fundamentais a ter em linha de conta passa por determinar:

- c) Até que ponto cada entidade beneficiará, do ponto de vista financeiro, da adesão a uma rede de preservação comum com custos partilhados, bem como;
- d) Qual o número mínimo de aderentes considerados necessários para tornar vantajosa a opção por uma solução partilhada.

Em todos os casos é necessário considerar variáveis como:

Dimensão total dos ODs a considerar;

Estimativa do seu crescimento anual;

Formatos dos ODs;

Sistemas e processos de armazenamento utilizados;

Requisitos de rapidez de acesso;

Número de cópias anualmente solicitadas a partir dos ODs;

Tempos de espera considerados adequados para a sua obtenção.

A análise para a definição do modelo de custos contemplou diferentes etapas e metodologias, a seguir enunciadas e acompanhadas dos respetivos resultados.

5.1. Pedido de orçamento para a contratualização dos serviços por parte de empresas especializadas

O objetivo era o de obter informação relativamente ao ponto a) atrás referido.

¹² Foram consultadas duas empresas: *Preservica*. Cf. sítio web. [Consult. 13 abril 2015]. Disponível em WWW: <URL: <http://preservica.com/> > e *Exlibris*. Cf. Sítio web. [Consult. 05 maio 2015]. Disponível em WWW: <URL: <http://www.exlibrisgroup.com/category/RosettaOverview> >.

¹³ DIREÇÃO-GERAL DO LIVRO, DOS ARQUIVOS E DAS BIBLIOTECAS - Concurso público: Caderno de encargos. Aquisição de serviços informáticos. Desenvolvimento de funcionalidades e *upgrade* do sistema RODA (Repositório de Objetos Digitais Autênticos). 2015.

Os orçamentos foram solicitados por dois dos elementos do GT, partindo do cenário atualmente existente nas entidades em que se encontram profissionalmente integrados. Contemplam, pois, custos de preservação digital suportados individualmente por uma organização.

Os resultados da estimativa orçamental foram os seguintes:

Empresa *Preservica*

Armazenamento em <i>cloud</i> (<i>Amazon</i>)	ca. 4.669 GB (4,56 TB)	Acesso imediato (informação disponível no momento)	20.346€/ano/GB
		Acesso diferido (informação disponível algumas horas após o envio do pedido)	15.466€/ano/GB

Uma simples conversão para TB e operando a correspondente operação matemática, permite obter o custo por TB/ano que importa em:

Armazenamento em <i>cloud</i>	Acesso imediato	4.462,26 € ano/TB
	Acesso diferido	3.391,99 € ano/TB

Foi possível calcular, a partir de informação fornecida por empresas, um custo médio de *storage* para *cloud* e para *datacenter* próprio. No primeiro caso o custo de **859,42 € TB/ano** corresponde à média de preços obtidos na *cloud* da *Amazon* e na *Azure* da *Microsoft*.

É possível então concluir que, no caso analisado, além do custo médio do *storage*, há uma diferença de **c. 3.605,84 €** na primeira modalidade e de **2.532,57 €** na segunda modalidade.

Admite-se que esta diferença constitua um acréscimo cobrado por serviços específicos de preservação digital. Este cálculo é relevante pois permite estabelecer uma fasquia no valor imputável a este tipo de serviços.

5.2. Cálculo das despesas implicadas na preservação digital, em diferentes tipos de cenário, para cada entidade - a título individual e de forma partilhada, com base numa rede comum

Os cálculos foram inicialmente efetuados com recurso a ferramentas disponíveis para o efeito, previamente identificadas e a seguir apresentadas.

Seguidamente foi elaborada uma matriz de cálculo própria, destinada a servir de referência comparativa para avaliar a pertinência dos resultados obtidos com recurso às ferramentas já mencionadas.

5.2.1 Levantamento dos modelos disponibilizados para o cálculo dos custos

Atualmente são disponibilizadas várias ferramentas e modelos cujo objetivo é o de facilitar o cálculo dos custos implicados na preservação digital. Uma vez identificadas, impôs-se perceber, através da sua análise, até que ponto cada uma delas se adapta e responde às necessidades específicas das diferentes entidades interessadas na preservação digital e aos diferentes tipos de cenários descritos.

O projeto *4C: Collaboration to Clarify the Costs of Curation*, financiado pela União Europeia no âmbito do 7.º programa quadro, que visa disponibilizar recursos que suportem o processo de gestão dos custos da preservação digital, procedeu a uma avaliação destes modelos, disponibilizando, no seu sítio web, as respetivas descrições e características¹⁴. Para o efeito, começou por construir uma terminologia de base a utilizar e por proceder a uma rápida caracterização dos respetivos custos e benefícios dos modelos. Indica como os modelos avaliados foram identificados e disponibiliza, para cada um, uma tabela com características, o que os torna facilmente comparáveis.

Os parâmetros com base nos quais a avaliação foi levada a cabo decorreram do levantamento das necessidades declaradas pelos *stakeholders* no que à informação de custos se reporta. Com base nelas foram apuradas as lacunas mais relevantes em cada modelo e feitas recomendações gerais.

Para informação mais detalhada sobre a metodologia adotada, cf. *D3.1—Evaluation of Cost Models and Needs & Gaps Analysis*¹⁵.

5.2.2 Seleção dos potenciais modelos a utilizar para a definição dos custos no âmbito da uma rede comum de preservação digital

Com base na avaliação efetuada foram selecionados os modelos que se afiguraram mais adequadas para dar resposta aos diferentes tipos de cenários a considerar:

- Cost Model for Digital Preservation (CMDP)¹⁶;
- LIFE3 Costing Model (LIFE3)¹⁷;
- NASA Cost Estimating Tool (NASA-CET)¹⁸.

¹⁴ Cf. Summary of Cost Models. In *4C: Collaboration to Clarify the Costs of Curation*. Sítio web. [Consult. 13 abril 2015]. Disponível em WWW: <URL: <http://4cproject.eu/summary-of-cost-models>>.

¹⁵ In *4C: Collaboration to Clarify the Costs of Curation*. Sítio web. [Consult. 13 abril 2015]. Disponível em WWW: <URL: <http://4cproject.eu/d3-1>>.

¹⁶ Informação retirada de WWW: <URL: <http://4cproject.eu/summary-of-cost-models/16-community-resources/outputs-and-deliverables/108-cost-model-for-digital-preservation-cmdp>>.

¹⁷ Informação retirada de WWW: <URL: <http://4cproject.eu/summary-of-cost-models/16-community-resources/outputs-and-deliverables/105-life3-costing-model-life3>>.

¹⁸ Informação retirada de WWW: <URL: <http://4cproject.eu/summary-of-cost-models/16-community-resources/outputs-and-deliverables/104-nasa-cost-estimating-tool-nasa-cet>>.

Através da sua utilização foi possível apurar custos individuais para cada potencial aderente, com base na informação reportada para as respetivas entidades pelos elementos do GT:

Modelo LIFE		
Objetos de formatos diversificados	período de 10 anos	605.000 euros
Imagens	período de 10 anos	806.000 euros

Modelo MODS		
Vídeo e imagens	período de 10 anos	2.796.901 euros
Todos os formatos (bases de dados, vídeos, imagens, texto, etc. - sem ligação com realidades reportadas pelos membros do GT)	período de 10 anos	13.527.578 euros

Modelo NASA-CET		
Não foi possível alcançar resultados		

Quadro 12 - Custos individuais a suportar pelos aderentes, apurados através da utilização dos modelos selecionados

No final desta abordagem, foi possível concluir:

- a) As ferramentas disponíveis estão orientadas para calcular os custos da preservação digital enquanto atividade.
- b) Não estão orientadas para o cálculo de custos de um repositório comum.

Tornou-se assim clara a necessidade de tentar um outro tipo de abordagem, até por se considerar relevante a existência de uma fonte independente para confirmar / infirmar os resultados individuais obtidos através dos referidos modelos.

5.2.3 Cenário alternativo

Foi elaborada uma matriz de cálculo, destinada a determinar uma unidade de custo base que permita, a cada potencial aderente, calcular os custos de adesão a uma rede comum. Esta matriz pode ser configurada em função do modelo de governação e da alocação de recursos humanos necessários à governação, gestão e manutenção da rede.

Esta matriz, que corresponde ao anexo 8.5 do presente relatório, contempla um conjunto de dados distribuídos da forma que abaixo se explicita.

Foram desde logo identificados dois tipos de custos:

Fixos	a repartir de forma equitativa pelo conjunto das entidades aderentes
Variáveis	cuja repartição será diretamente proporcional ao grau de utilização dos serviços solicitados à rede por cada entidade aderente

Para o caso em análise foram identificadas outras duas categorias de custos, mas que não influenciam o método seguido:

Custos operacionais e de investimento (OPEX e CAPEX)

Custos diretos e indiretos

Neste cenário apenas foram considerados custos diretos.

Constata-se que a maior parte dos custos são fixos, correspondendo a despesas com:

Recursos humanos

Custo de manutenção e atualização do repositório

Custos de manutenção de edifício (espaço físico ocupado, eletricidade, segurança, etc.)

Outros custos (computadores para utilização administrativa, material de escritório, etc.)

Tendo em conta que uma das exigências para a certificação de repositórios de preservação digital passa pela existência de recursos humanos afetos em exclusividade à gestão do repositório, optou-se por um cenário estático em que há um conjunto de técnicos a trabalhar no repositório a tempo inteiro. Inclui fundamentalmente o valor dos salários.

De referir que as duas últimas parcelas assumem um papel residual em relação às despesas com recursos humanos e com custos de manutenção e atualização do repositório.

Poder-se-ia optar, caso se pretendesse transformar custos fixos em variáveis, de forma a diminuir o ónus financeiro permanente, por indexar o número de recursos humanos afetos ao repositório ao volume de trabalho exigido pela rede comum, expresso com base numa dupla vertente:

Número de aderentes (mais trabalho de natureza administrativa)

Número de ODs incorporados / a incorporar (mais trabalho de natureza técnica)

No entanto este cenário implicaria dois pressupostos:

1/ A possibilidade de adquirir e dispensar recursos humanos através de um esquema de contratação de trabalho flexível, ou seja, a possibilidade de contratar a prazo segundo picos identificados de trabalho;

2/ A afetação dos recursos humanos disponíveis a outros trabalhos não diretamente ligados a preservação digital. Neste caso assume-se uma estrutura com carga laboral não exclusivamente dedicada ao repositório.

As tarefas a desempenhar - específicas de preservação digital e administrativas - foram retiradas das ferramentas acima descritas:

	Gestão do repositório
1	Planeamento
2	Monitorização de tecnologia
3	Negociação
4	Atividades gestionárias
5	Atividades de monitorização e controlo
6	Aprovisionamento e economato
7	Contabilidade
8	Instalações
9	Ingestão
10	Monitorização
11	Acesso
12	Edição de metadados
13	Migração
14	Manutenção de plataforma tecnológica
15	Gestão de utilizadores
16	Prestação de serviços
17	Desenvolvimento evolutivo do repositório

Quadro 13 - Tarefas a desempenhar na gestão do repositório

O primeiro cenário, no contexto português e da Administração Pública, não parece muito viável, enquanto o segundo não se afigura compaginável com as exigências de certificação

de repositórios digitais patentes na norma ISO 16363:2012¹⁹. Esta certificação é um requisito de implementação aparentemente necessária, se levarmos em conta os resultados do inquérito público realizado pelo grupo de trabalho que aponta para uma perceção muito valorizada da salvaguarda da autenticidade da informação custodiada²⁰.

Assumiram-se, portanto, os custos de pessoal como **custos fixos**, sendo o número de pessoas a afetar à gestão do repositório calculado para um período de **10 anos**.

Foram utilizados dados salariais das tabelas de vencimentos disponibilizadas pela Direção-Geral do Livro, dos Arquivos e das Bibliotecas. Estes são muito variáveis, pois oscilam, para cada carreira, numa longa escala composta por diferentes graus e posições. Optou-se por seleccionar valores intermédios, pese embora se pudesse ter recorrido à média de todos os valores salariais. No entanto, a eliminação dos *outliers* seria necessária, de forma a obter um resultado estatisticamente relevante. Esse esforço não foi, no entanto, realizado para este relatório.

Carreira	Grau e posição	Salário/ano	Salário/mês
Técnico superior	grau 8 posição 39	34.122,06 €	2.437,29 €
Assistente técnico	grau 7 posição 12	14.658,00 €	1.047,00 €
Chefia	intermédia grau 1	36.802,92 €	2.628,78 €
Informático especialista	espc. grau 3 nível 2 pos. 860	41.330,94 €	2.952,21 €
Informático técnico	tec. grau 2 nível 1 pos.680	32.680,20 €	2.334,30 €
Diretor	superior grau 1	62.749,26 €	4.482,09 €

Quadro 14 - Dados salariais relativos às diferentes carreiras

Enquanto **custo variável**, há a considerar o espaço de armazenamento ocupado ou a ocupar pelo repositório, que corresponderá ao somatório das necessidades de armazenamento de cada uma das entidades aderentes, pelo que se trata de um custo individualizável. Trata-se também de um custo potencialmente evolutivo, proporcional às incorporações adicionais eventualmente realizadas no repositório pelos aderentes.

No entanto, por si só, o espaço de armazenamento não basta para obter o custo real. Há ainda que ter em conta a complexidade do processamento necessário para realizar serviços de preservação digital. Esta métrica está diretamente relacionada com o número dos ODs a

¹⁹ ISO 16363:2012 - Space data and information transfer systems - Audit and certification of trustworthy digital repositories. Genève: International Standard Organization.

²⁰ Ver Ponto 4 do presente relatório, mais especificamente as conclusões relacionadas com a importância concedida pelos respondentes às propriedades que determinam a autenticidade dos ODs.

processar e respetivo esforço de computação. Assim, a obtenção de uma unidade base de custo passa pela obtenção de um rácio entre as seguintes duas variáveis:

Dimensão dos objetos expressa numa unidade binária ²¹ (MB, TB, etc.)	x	(número de objetos que representam essa dimensão x custo de processamento)
---	---	--

A título de exemplo, observe-se o seguinte caso:

Diferentes possibilidades de ocupação de 1 TB de armazenamento	
2 ODs	500 GB cada
1.000 ODs	1 GB cada

Em teoria o segundo caso seria mais dispendioso, pois o esforço de processamento é tanto maior (podemos considerá-lo diretamente proporcional) quanto maior o número de objetos a processar.

As simulações efetuadas com o material e valores declarados por alguns membros do grupo de trabalho permitiram demonstrar que os custos seriam diferentes caso optássemos por apenas um dos critérios - critério de dimensão ou pelo critério de número de ODs - para calcular o custo variável:

DGLAB	15M de objetos para 300 TB de espaço útil ocupado
SIC	c. 74K objetos para 2 PB de espaço útil

No caso da DGLAB, por exemplo, os valores seriam consideravelmente superiores aos da SIC. No entanto, usando o critério dimensão, verifica-se precisamente o contrário. Parece portanto necessário considerar algumas questões:

Questão 1: encontrar um rácio equilibrado entre estas duas variáveis:

- dimensão dos objetos expressa numa unidade binária²² (MB, TB, etc.)
- número de objetos que representam essa dimensão.

Não foi ainda possível apurar se uma relação linear será a mais adequada, considerando a disparidade das escalas de valores em questão.

²¹ Por uma questão de uniformização de medida, sugere-se a utilização de uma única unidade binária: o TB, podendo recorrer-se a submúltiplos quando necessário: por ex. um GB será equivalente a 0.000976 TB.

²² Por uma questão de uniformização de medida, sugere-se a utilização de uma única unidade binária: o TB, podendo recorrer-se a submúltiplos quando necessário: por ex. um GB será equivalente a 0.000976 TB.

Importa ainda saber como contabilizar a métrica ou esforço subjacente à manipulação de ODs. Este valor é essencial, assumindo que é ilógico considerar apenas o número de ODs como parcela da equação para obter o custo base. Com efeito, para obtermos um custo é necessário multiplicar o número de ODs por uma grandeza que tem de ser determinada.

Assim, o problema seguinte consiste em:

Questão 2 determinar as variáveis possíveis e adequadas para estimar o custo de computação, obtendo-se assim o fator a relacionar com o número de ODs

Hipótese 1: Neste contexto considerámos como hipótese possível determinar o esforço computacional exigido para o efeito e converter essa grandeza num custo mensurável em unidades monetárias. O problema reside, no entanto, em selecionar métricas aceitáveis para calcular esse esforço: uma possibilidade equacionada é o desempenho computacional (ver [WWW <URL: http://en.wikipedia.org/wiki/Computer_performance>](http://en.wikipedia.org/wiki/Computer_performance)) que apresenta métricas em princípio coadunáveis com o nosso propósito. No entanto a complexidade subjacente ao cálculo bem como a aparente escassez de exemplos de aplicação práticas destas métricas levaram-nos a suspender, pelo menos provisoriamente, este caminho.

Hipótese 2: Outra hipótese ainda é a métrica designada de eficiência algorítmica, não tendo havido ainda possibilidade de a explorar mais detidamente. [WWW <URL: http://en.wikipedia.org/wiki/Algorithmic_efficiency#Measures_of_resource_usage >](http://en.wikipedia.org/wiki/Algorithmic_efficiency#Measures_of_resource_usage)

De referir que o custo de computação cai exponencialmente de ano para ano. Assim, e tendo em conta este aspeto, optou-se por considerar um cenário em que se cobraria aos aderentes em função do **espaço de memória alocável**, proporcional ao número de ODs declarados. Em princípio, a memória implica uma maior despesa, já que, embora leve mais tempo a processar objetos maiores do que mais pequenos, o custo computacional não só é pequeno, com revela tendência decrescente, ano após ano.

O custo do espaço de memória é calculado com base no custo dos GB (normalmente a dimensão da memória de um computador ou servidor é medida na ordem dos GB e não dos TB).

A fórmula final para obter o custo base poderá assumir uma configuração geral do tipo:

$$\text{Unidade Base Custo} = \text{Custos Fixos} \times \text{Custos Variáveis}$$

$$\text{Custos Variáveis} = (\text{dimensão_espaço} \times \text{n.º_ODs}) (\text{esforço_computação})$$

Note-se que caso os aderentes não pretendam obter o serviço de armazenamento, por já disporem eventualmente dessa estrutura e não quererem replicar custos, será retirada da equação a parcela correspondente, havendo sempre lugar ao pagamento da parcela **“número de ODs x esforço computação”**.

Relativamente ao custo da memória é possível determiná-lo a partir do total da memória alocada a um servidor e dividindo esse valor por 3.

O valor 3 corresponde ao número de anos que, em média, se consideram necessários para amortizar o custo de um computador / servidor. Desta forma obtemos o valor anual de utilização de memória por servidor. A seguir, o valor assim obtido deverá ser multiplicado pelo número total de servidores utilizados pela estrutura tecnológica comum para processar os ODs ingeridos.

Exemplo: Se tivermos, em média, um servidor com 10 GB de memória, dividimos 10/3 (anos), obtendo o valor anual de um servidor equivalente a 3,33 GB utilizados por ano. Se o parque de servidores alocados ao serviço de preservação digital for de 10, teremos então:

$$3,33 \times 10 = 33,3 \text{ GB de memória despendidos no serviço preservação digital}$$

Para percebermos qual o valor de memória necessária por objeto digital é necessário dividir o valor total de memória anual pelo total de ODs ingeridos no repositório:

$$\text{Para 15M de objetos:} \quad 33,3 \text{ (GB)} / 15.000.000 \text{ (objetos)} = 0,00000222$$

cada OD consome 0,00000222 GB de memória

Para obtermos, seguidamente, o custo imputável a cada aderente, haverá que multiplicar o valor unitário de memória/objeto digital obtido, pelo número de objetos que cada aderente ingeriu ou pretende ingerir no repositório.

Imaginemos que dos 15M referidos, 1M é do produtor/aderente A - iremos multiplicar o valor unitário memória/objeto digital por 1M, ou seja:

$$\text{Aderente A} \quad 1\text{M (num total de 15 M):}$$

$$0,00000222 \times 1.000.000 = 2,22 \text{ GB}$$

Este valor corresponde à memória solicitada, expressa em GB (a nossa unidade de medida para a memória) para gerir 1M de objetos, i.e., 2,22 GB de memória.

Agora é apenas necessário multiplicar este valor pelo custo estimado do GB de memória, para obter o custo efetivo imputável ao aderente A no que respeita a memória.

5.2.3.1 Decrementação da unidade base de custo

O último aspeto que foi tido em conta consistiu em determinar qual o grau de decrementação dos custos em função do aumento de número de aderentes.

Considerando que a rede terá características cooperativas e de partilha de custos, em princípio quantos mais aderentes menores os custos. Com efeito, a adesão de um novo aderente representa uma unidade mais que suporta os custos fixos, já que os custos variáveis irão depender exclusivamente do material que cada aderente queira ingerir no repositório.

Assim, considerámos as seguintes possibilidades, em que os custos fixos são arbitrários:

1. 5 aderentes criam um serviço que custa 100.000 euros por ano de custos "fixos" e cada um paga ainda de armazenamento/processamento dos ODs 50.000 euros. Os 100.000 dividem por 5 = 20.000, donde cada paga 70.000 no primeiro ano.
2. Ao fim de 1 ano aparece outro aderente. Isso significa que a partir do segundo ano há 6 entidades a dividir todos os custos, por isso temos que cada um dos 6 aderentes irá agora pagar 100.000 a dividir por 6 (16.667) mais os 50.000 do armazenamento/processamento, ou seja, 66.667.

Cenário 1

Os "fundadores" têm o benefício de haver mais um aderente, mas não são recompensados por terem tido a iniciativa de aderir cedo, viabilizando o serviço. Por outro lado, o novo aderente não teve o benefício do sistema durante 1 ano, portanto também não é justo cobrar-lhe um ano "a posteriori" de um serviço que não usou.

Ano 1	5 aderentes	pagam 20.000 + 50.000
Ano 2	5 aderentes + 1 novo aderente	pagam 16.557 + 50.000

Neste caso os fundadores têm vantagem em atrair mais aderentes para a rede, pois tal significa menos custos. A vantagem não é retroativa, verificando-se apenas a partir da adesão de novos aderentes.

Cenário 2

Mantém-se a lógica da variação 1 mas inclui-se o pagamento de uma "jóia de inscrição", destinada a alimentar um fundo de reserva da rede para aquisições de material, equipamentos, serviços, etc.

Conclusões

5.3 Os custos de adesão a uma estrutura de preservação digital comum a suportar pelos aderentes, segundo o que foi possível apurar, são eventualmente excessivos.

5.4 Face a essa constatação, concluiu-se que será necessário equacionar diversas formas de financiamento da rede, por forma a garantir a sua independência. Para o efeito, e para além das despesas de preservação a suportar pelos aderentes, poderão existir outros serviços igualmente pagos - monitorização das tecnologias constitui um exemplo. Poder-se-á ainda considerar o mecenato ou as doações, mas poderão ser equacionadas outras alternativas.

5.5 Ao concluir a análise e a posterior proposta relativas ao modelo de custos, um dos principais aspetos a considerar prende-se com o facto de o GT não dispor de recursos dotados de competências e experiência na área financeira, e não ter conseguido obter a colaboração de profissionais especializados na elaboração do modelo de custos. Assim, a proposta apresentada, sendo embora a possível, carece de maior consistência, evidenciando aspetos que necessitam de maior aprofundamento e clarificação, bem como do contributo de profissionais especializados.

5.6 Há ainda a mencionar que o modelo de custos é inseparável do tipo de modelo de governação a adotar, já que os custos são diretamente influenciados pelo tipo de gestão adotado para a rede. Uma maior ou menor participação de aderentes nas tarefas de governação, gestão e manutenção da rede, dispensando ou obrigando à contratualização externa dos recursos humanos necessários para as levar a cabo, determinará uma redução ou um aumento dos custos, podendo por isso contribuir para viabilizar ou inviabilizar a sua sustentabilidade.

5.7 Este aspeto assume uma relevância tanto maior quanto se constatou que os recursos humanos necessários à governação, gestão e manutenção da rede, bem como o armazenamento dos ODs a preservar no repositório, correspondem às componentes mais caras.

5.8 Relativamente ao armazenamento, foi possível concluir que quanto maior for a necessidade de memória por parte de cada aderente, maior será a diferença entre o valor que terá que suportar, consoante pretenda ou não contratualizar armazenamento fornecido pela rede. Será necessário prever as situações em que os aderentes já possuam um serviço desta natureza, próprio ou contratualizado. Nestas circunstâncias a opção por não recorrer a um serviço comum de armazenamento tem de ser considerada, deduzindo-se os custos respetivos do compromisso a assumir pelo potencial aderente.

5.9 Concluiu-se igualmente que as formas de contornar os elevados custos de preservação digital carecem de investigação e desenvolvimento adicionais, tendo sido avançada, como uma possibilidade de redução de custos, a preservação digital distribuída, em que cada aderente disporia das suas próprias soluções de armazenamento, cabendo ao repositório de preservação digital aceder remotamente ao repositório de cada entidade aderente para levar a cabo as necessárias ações de preservação digital. Este tipo de solução teria a vantagem de reduzir os custos no âmbito de uma rede partilhada.

5.10 Um outro aspeto a ter em conta é a dificuldade de calcular com precisão aspetos como o crescimento das necessidades das entidades aderentes, ou seja, os custos variáveis. Enquanto os custos fixos terão tendência a decrescer face ao aumento do número de entidades aderentes, os custos variáveis podem ser tendencialmente imprevisíveis.

5.11 Do inquérito realizado no âmbito do projeto, relativamente à disponibilidade das instituições para contribuírem financeiramente para um repositório comum de preservação digital, como contrapartida de serviços prestados (taxa de resposta: 27,4%), a maior parte dos respondentes discorda ou discorda inteiramente, como aliás se pode constatar através da análise do gráfico que se segue:

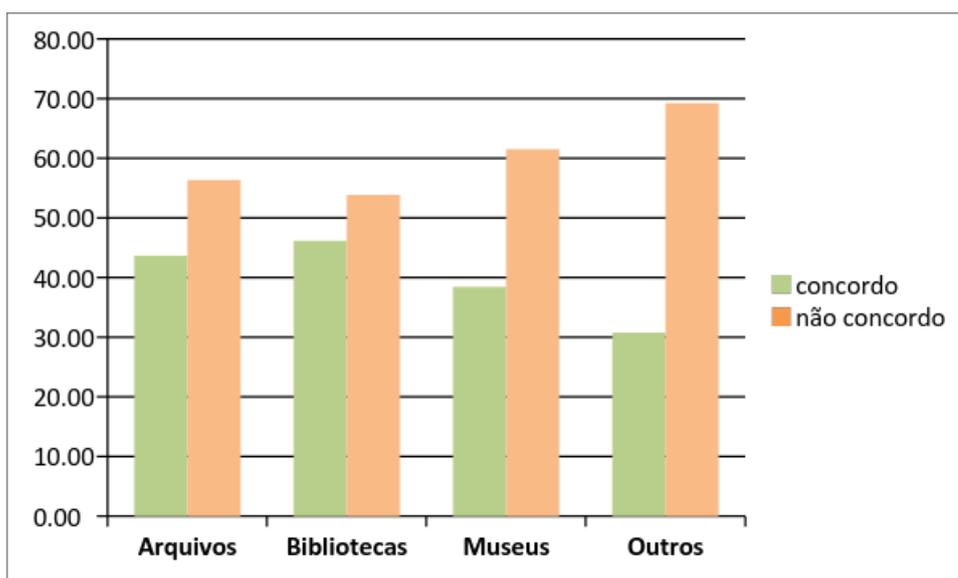


Gráfico 6 - Disponibilidade para contribuir financeiramente para um repositório comum como contrapartida de serviços prestados

Este resultado ficará provavelmente a dever-se ao facto de a maior parte das instituições não ter disponibilidade para suportar as despesas inerentes à adesão a uma rede partilhada, aspeto que parece sair reforçado face à esmagadora percentagem de concordância, por parte dos respondentes (taxa de resposta: 29,2%), de que gerir património digital em comum com outras instituições tem vantagens financeiras, como se pode constatar através da análise do gráfico que se segue. De salientar a quase total ausência de respostas que apontam para a total discordância e para a fraca expressividade das que indicam discordância.

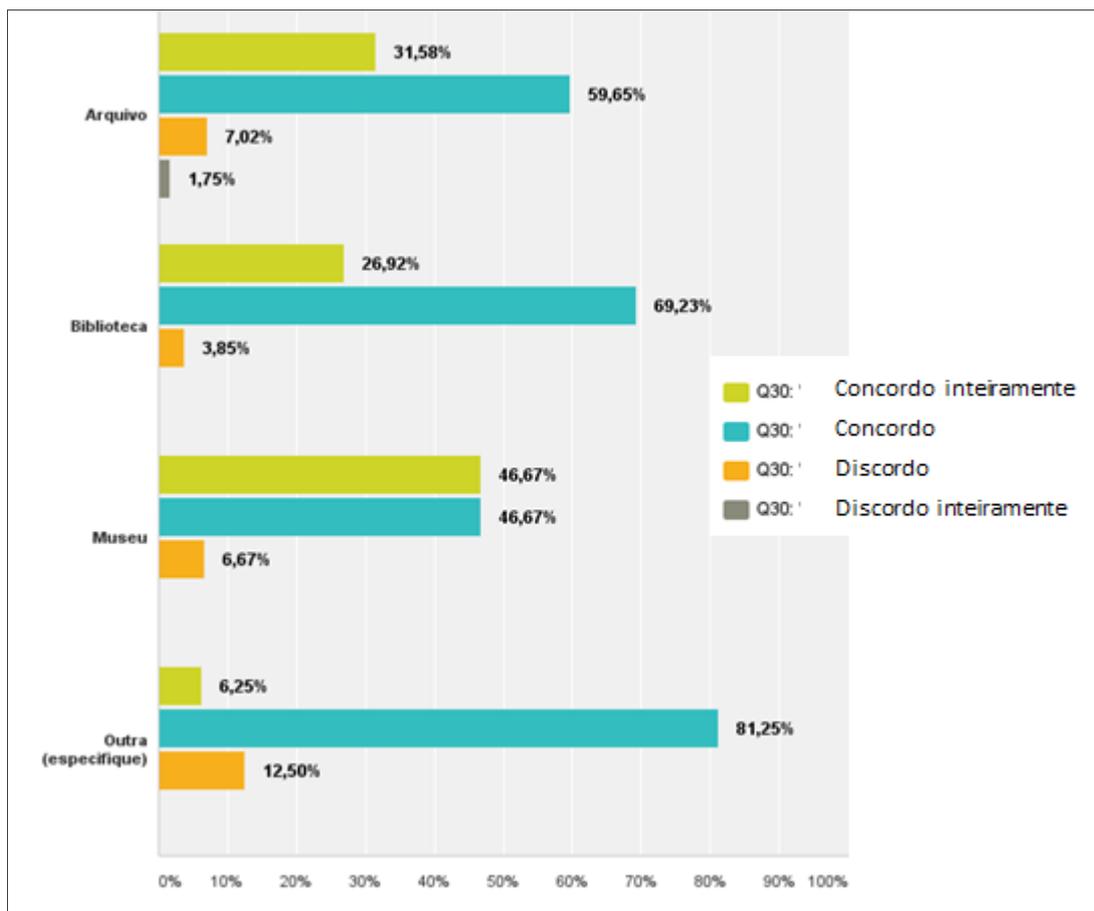


Gráfico 7 - Graus de concordância/discordância das CdP face à eventual disponibilidade para contribuir financeiramente para um repositório comum

5.12 Uma das hipóteses alternativas de financiamento da rede poderia passar pela criação de um modelo de exploração comercial, com cedência de uma percentagem sobre a venda de cópias de objetos digitais. A resposta dos respondentes à questão que equaciona esta possibilidade (taxa de resposta: 25,7%), como pode constatar-se através da análise do gráfico, as opiniões dividem-se: Museus e Arquivos concordam, embora a percentagem de respostas positivas seja muito mais expressiva nos Museus; Bibliotecas e as demais CdP discordam, sendo a percentagem de discordâncias muito mais expressiva nas outras CdP.

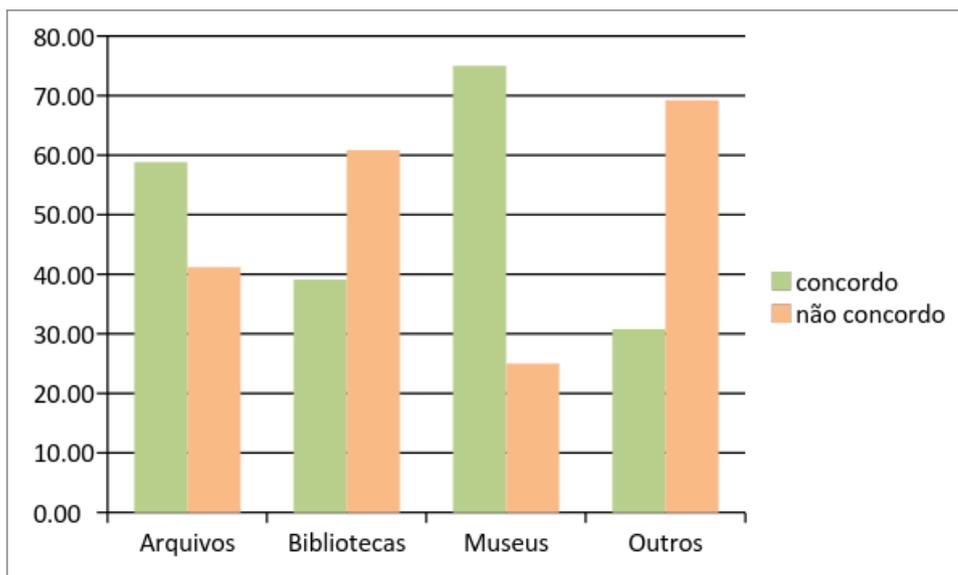


Gráfico 8 - Disponibilidade para a implementação de um modelo de exploração comercial, associado ao repositório, com cedência de uma percentagem sobre a venda de cópias de ODs

5.13 As respostas às questões relacionadas com o modelo de custos, no âmbito do inquérito realizado, reforçam as conclusões do GT, e apontam para o facto de os custos de adesão a uma rede de preservação digital comum a suportar pelos aderentes, segundo o que foi possível apurar, são excessivos e de que será necessário equacionar outras formas possíveis de financiamento da rede que impliquem um esforço menor por parte dos aderentes.

6 Modelo de governação

A definição de um modelo de governação implica a aferição das características, dos princípios, do objetivo, da missão e das atividades da rede, bem como do tipo de governação a adotar e dos diferentes órgãos a considerar para a materializar. A governação da rede está diretamente relacionada, como aliás foi referido no ponto 5 do presente relatório, com o modelo de custos a adotar. Admitiu-se desde o início ser necessário procurar fontes de financiamento próprias que, em parte ou no todo, deverão assentar em contributos estipulados para os aderentes.

Ao longo deste processo o GT concluiu que embora haja vários cenários possíveis para um modelo de governação, há um conjunto de funções comuns a qualquer ação gestionária que deverão ser consideradas. No entanto, a definição específica dos órgãos ou a articulação entre os mesmos são objeto de influência de muitas variáveis pelo que optou por não os incluir neste cenário. Considerou-se ainda dever este tópico ser objeto de investigação mais aprofundada. Foram igualmente identificados obstáculos de natureza jurídica e orçamental que implicam maior profundidade e averiguação. O GT elaborou um modelo possível de rede cujas características, a seguir explicitadas, foram escrutinadas através do segundo inquérito

realizado de forma a obter-se a perceção generalizada por parte das CdP face a essas mesmas características e ao conceito de rede partilhada.

Neste contexto, apresentam-se seguidamente os princípios propostos, que assumem como pressupostos uma rede participativa e não hierarquizada.

6.1 Princípios

- Adesão voluntária e livre. Abertura da rede a todas as pessoas singulares ou coletivas, com interesse em utilizar os serviços por ela disponibilizados e dispostas a assumir as responsabilidades de membro.
- Gestão participativa. Rede a gerir pelos membros, os quais participam ativamente na formulação da sua política e opções estratégicas e na tomada de decisões, com iguais direitos de votos.
- Participação económica dos membros. Estes contribuem de forma proporcionada - de acordo com os serviços que subscreverem - para a sustentabilidade da rede, assumindo o seu controlo económico.
- Autonomia e independência. A rede pretende operar de forma autónoma e independente, não obstante o estabelecimento eventual de acordos de cooperação com outras entidades, tendo em vista a superação de lacunas funcionais e o potenciar de conhecimento.
- Formação e informação. No âmbito da prossecução da sua missão, a rede promove a informação e formação dos seus membros através de diversas iniciativas, podendo estas ser também extensivas a outras entidades e público em geral.
- Intercooperação. A fim de potenciar a eficácia da sua ação a nível nacional, a rede pode cooperar com entidades pertencentes a diferentes níveis de administração, designadamente local, distrital, regional e internacional.
- Interesse pela comunidade. Manifestado nos propósitos de:
 - Recolher e salvaguardar informação digital a que seja reconhecido valor patrimonial.
 - Viabilizar, através das respetivas entidades detentoras, o acesso e usufruto pelo público desse património.
 - Transmitir às gerações vindouras o legado de património digital nacional, devidamente preservado.
- Princípio da neutralidade. A rede não visa substituir-se às entidades aderentes na responsabilização pela preservação dos seus ODs. Embora a ingestão de ODs no repositório obedeça a um conjunto mínimo de requisitos, cabe às entidades aderentes a avaliação de quais, de entre os ODs por elas produzidos ou custodiados, os que detêm valor patrimonial e, enquanto tal, devem integrar um repositório partilhado de preservação digital. Cabe-lhes igualmente assegurar que os direitos que recaem sobre a informação neles contida - patrimoniais, de autor e conexos - é respeitada.

6.2 Organização

A organização da rede baseia-se num modelo estático. Embora o GT tenha desenvolvido análise bastante mais extensa relativamente a aspetos de relações entre entidades e

comportamentos gestionários possíveis, considerou-se que essa análise carece de estudo e fundamentação mais aprofundados, pelo que não serão apresentadas no presente relatório. (ver secção 7 do presente relatório).

6.2.1. Âmbito

As características propostas para a rede no que se refere ao seu âmbito são as seguintes:

- **Nacional.** Não se exclui, no entanto, a possibilidade de abertura e a colaboração com entidades estrangeiras;
- **Aberta** a todas as instituições públicas ou privadas que detenham património digital e que possuam a vontade de empreender ações concretas e comprometidas no sentido de o preservar prolongadamente;
- **Transversal**, porque pretende abranger os mais diversificados e variados domínios culturais e CdP;
- **Descentralizada**, porque todos os aderentes podem participar na sua governação e na definição das políticas e prioridades a adotar.

6.2.2. Objetivo e missão

O propósito essencial da rede é a gestão participada de uma infraestrutura de preservação tendo em vista assegurar a sua sustentabilidade ao longo do tempo, promovendo a preservação de património digital.

A rede tem como missão promover a partilha de conhecimento, bem como a cooperação na definição de políticas, estratégias e ações visando a gestão, manutenção e atualização de um repositório partilhado, que assegure a preservação e o acesso continuado aos ODs patrimoniais das entidades aderentes de uma forma efetiva e sustentável.

A rede visa encorajar as entidades a responsabilizar-se e a organizar-se para assegurar a preservação dos seus ODs, mantendo sobre estes uma tutela ativa.

Visa ainda a divulgação, formação, promoção de boas práticas, a investigação e a inovação na área da preservação digital.

6.2.3. Serviços

Tendo em conta o desenvolvimento da rede e atendendo à concretização do seu objetivo e missão, as principais atividades a levar a cabo distribuem-se pelos seguintes domínios:

- Preservação digital;
- Divulgação para promoção da rede (realização de eventos, publicidade / marketing);
- Inovação (atividades que proporcionem acréscimo de valor à rede, resultantes da criatividade dos membros produtores);
- Formação (realização de *workshops*, cursos, seminários, publicações);
- Outras (que, entretanto, possam vir a ter cabimento e pertinência no âmbito da continuidade do desenvolvimento da rede).

Para a concretização destas atividades deverá ser suportado um conjunto de serviços:

#	Serviços
1	Ingestão e preservação de ODs;
2	Prestação de diferentes níveis de serviço relativamente à ingestão e gestão de ODs para cumprimento de política de gestão de documentos (exemplos: formatos cuja preservação não pode ser assegurada, servindo a sua ingestão apenas para efeitos de <i>backup</i> ; preservação temporária de ODs com mero valor administrativo, sujeitos a prazos de conservação, findos os quais os ODs serão eliminados, etc.);
3	Produção e divulgação de documentos orientadores e de boas práticas na área da preservação digital;
4	Monitorização de tecnologias (a ser desenvolvido no âmbito do Roda +);
5	Manutenção de uma plataforma tecnológica comum, baseada no modelo OAIS, compreendendo um núcleo central de funções de preservação e, eventualmente, disponibilização de serviço de armazenamento partilhado.

Quadro 14 - Serviços a suportar pela rede

6.3. Aderentes

OS aderentes poderão ser pessoas coletivas, públicas ou privadas, ou pessoas singulares. Terão direitos e deveres:

Direitos

- Participar nas atividades de gestão e de governo da rede, integrando os respetivos órgãos de governação, colaborando na definição da política e estratégia da rede, bem como na elaboração de planos de atividades, contribuindo ainda para apoiar a tomada de decisão, através do seu voto;
- Usufruir de todos os serviços disponibilizados no âmbito da rede;
- Proceder ao cancelamento da adesão.

Deveres

- Pagar as despesas e quotizações estipuladas de acordo com as modalidades de pagamento definidas;
- Assumir responsabilidade partilhada pela preservação do património digital ao cuidado da rede;
- Cumprir os requisitos técnicos e administrativos de adesão e ingestão de informação digital no repositório;
- Assegurar os desenvolvimentos necessários para articulação dos seus sistemas com o repositório.

É possível colocar a possibilidade de, à semelhança de experiências similares²³, considerar diferentes tipos de aderentes, determinados em função dos serviços pretendidos ou do grau de compromisso desejado.

²³ Ver por exemplo o Meta archive: WWW: <URL: <http://www.metaarchive.org/> >.

A responsabilidade pela preservação digital é partilhada, pelo que não cessa com a transferência de ODs para o repositório. Todos os membros aderentes terão de assumir responsabilidade relativamente à política de preservação digital adotada e aos ODs ao cuidado da rede (cf. ponto 6.1 “Princípios a adotar pela rede”).

A adesão à rede implica o pagamento de uma quotização anual por parte da totalidade das entidades aderentes. O objetivo é o de garantir a sua independência, sustentabilidade e estabilidade.

6.4. Governança

Foram identificadas funções comuns a qualquer tipo de organização e que correspondem às categorias de processos indispensáveis para decidir e executar. Uma rede comum não pode deixar de ter:

- Uma função deliberativa, exercida normalmente pelo conjunto de aderentes;
- Uma função executiva, exercida por um núcleo encarregue de executar e gerir a estrutura. Existem várias possibilidades para o exercício desta função, todas elas dependentes de variáveis que carecem de maior aprofundamento, por ex.:
 - Várias entidades aderentes;
 - Um subconjunto de aderentes fixos e outro rotativo;
 - Uma entidade fixa, definida por lei;
 - Uma instituição pública ou privada contratada para o efeito.
- Uma função tecnológica destinada à manutenção preventiva, corretiva e evolutiva do repositório comum;
- Um repositório tecnológico comum compreendendo a valência de receber e preservar objetos digitais;
- Uma função de investigação com o propósito de criar conhecimento na área da preservação digital que possa ser replicado e reutilizado pelos diversos aderentes da rede. Esta função poderia ser assegurada por grupos de trabalho temáticos que operem em conjugação com o sistema nacional de inovação composto pelas universidades e pelo tecido empresarial.

Associado ao repositório, poderá ser eventualmente considerada a existência de um modelo de exploração comercial, baseado na cobrança, às entidades aderentes, de uma percentagem fixa dos lucros auferidos com a venda de cópias dos ODs preservados (sobre a opinião dos respondentes ao inquérito a respeito desta possibilidade, cf. ponto 5.12 do presente relatório). Estas receitas destinar-se-iam a assegurar a manutenção e os desenvolvimentos adicionais requeridos pelo repositório. Caso um aderente optasse por esta modalidade, poderia cumprir total ou parcialmente a sua quotização recorrendo a este método. Qualquer diferença entre a receita obtida por este processo seria vertida para a rede e o montante total da quotização devida seria quitada pelo aderente.

Visando igualmente a manutenção e os desenvolvimentos adicionais do repositório, poderá ainda ser considerada a possibilidade de recurso a financiamentos alternativos, eventualmente publicidade.

O diagrama a seguir apresentado denota um esquema estrutural possível:

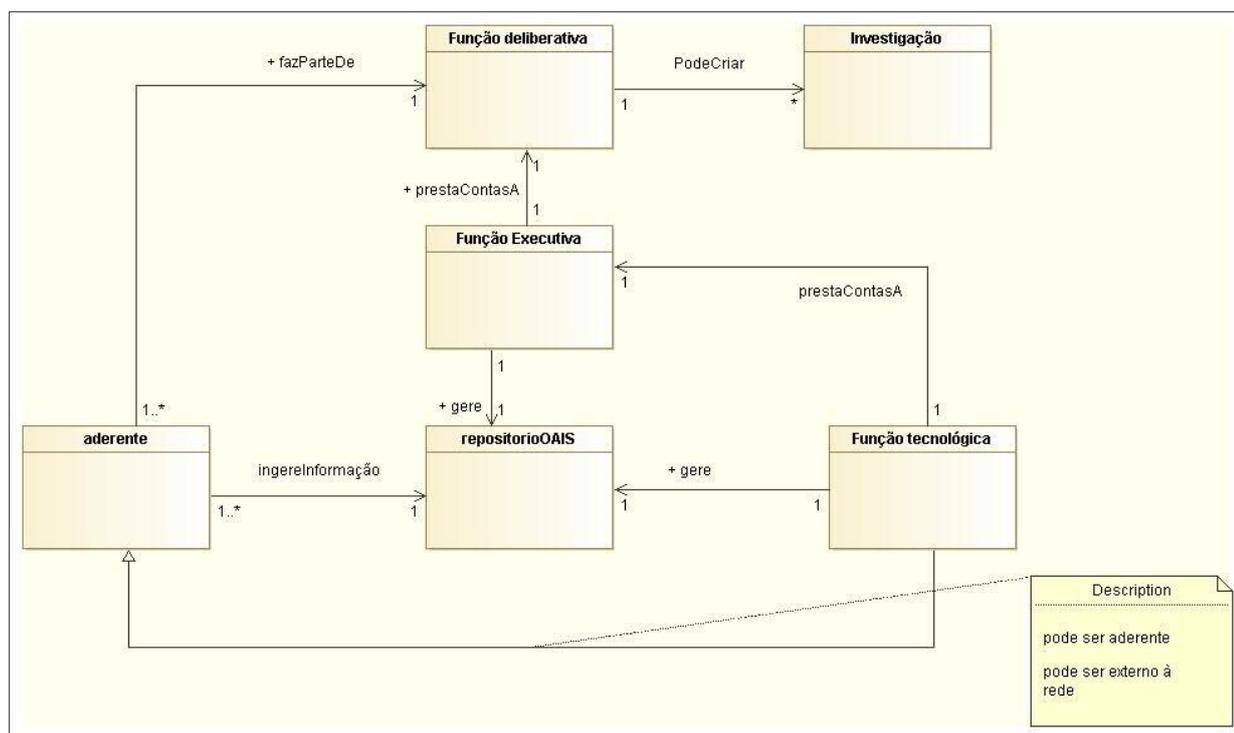


Diagrama 1 - Eventual esquema estrutural a adotar

Conclusões

A proposta apresentada carece de mais aprofundamento e de clarificação relativa a questões de natureza jurídica, designadamente, a resolução das situações administrativas referentes à afetação de recursos humanos e à formalização legal da rede.

O inquérito público levado a cabo no âmbito do projeto Continuidade Digital permite concluir, quanto às questões de governação da rede, o seguinte:

Todas as CdP consideram aceitável a participação das instituições na gestão coletiva de um repositório partilhado, convergindo maioritariamente na variável “concordo”, com valores percentuais bastante significativos (oscilam entre os c. 49% e os 68%), aceitando, portanto, cooperar com outras organizações a fim de **preservar, de forma comum, o património digital.**

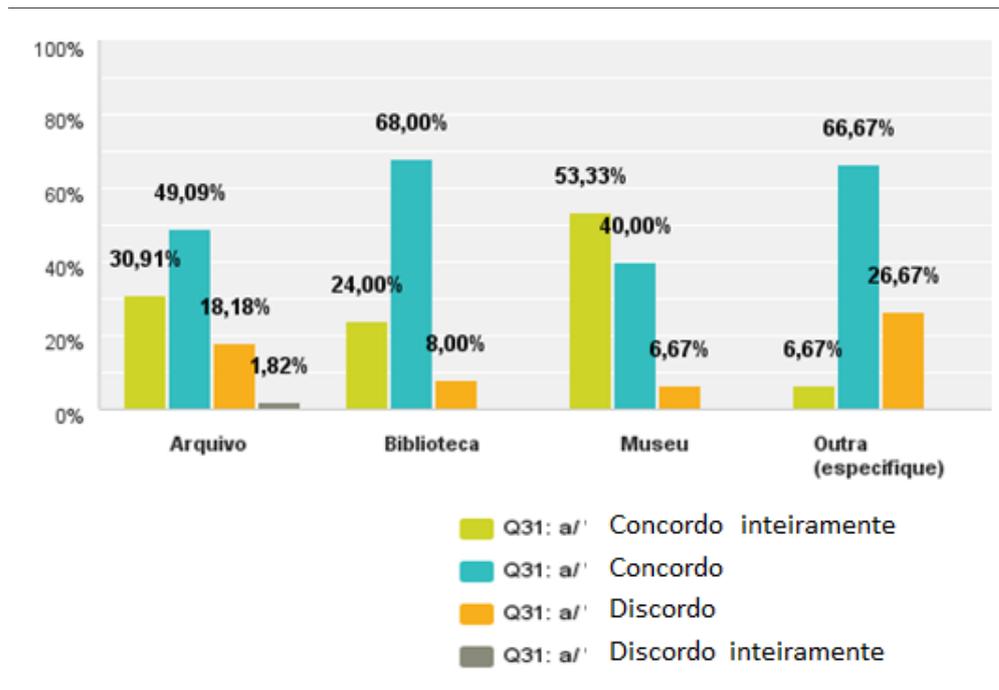


Gráfico 9 - Possibilidade de opção por um modelo descentralizado de gestão da rede

A maioria das instituições com custódia de património digital possivelmente não tem disponibilidade financeira para contribuir com uma quotização para o financiamento da RCPD (ver capítulo modelo de custos). Não obstante, ainda existe uma percentagem relativamente significativa - c. de mais de 30% - de instituições que admitem ter disponibilidade para contribuir para o financiamento da RCPD.

A maioria dos respondentes das várias CdP relativamente à exigência de **confiabilidade da entidade prestadora de serviços de preservação digital** opta preferencialmente por uma entidade pública. No entanto, algumas CdP, sobretudo, **Museu** e **Outra**, admitem também, poder confiar os seus ODs igualmente a entidades privadas, desde que consideradas confiáveis, condicionando este resultado o modelo de governação por ter de considerar as duas hipóteses: uma gestão tecnológica por um parceiro ou por uma entidade privada externa.

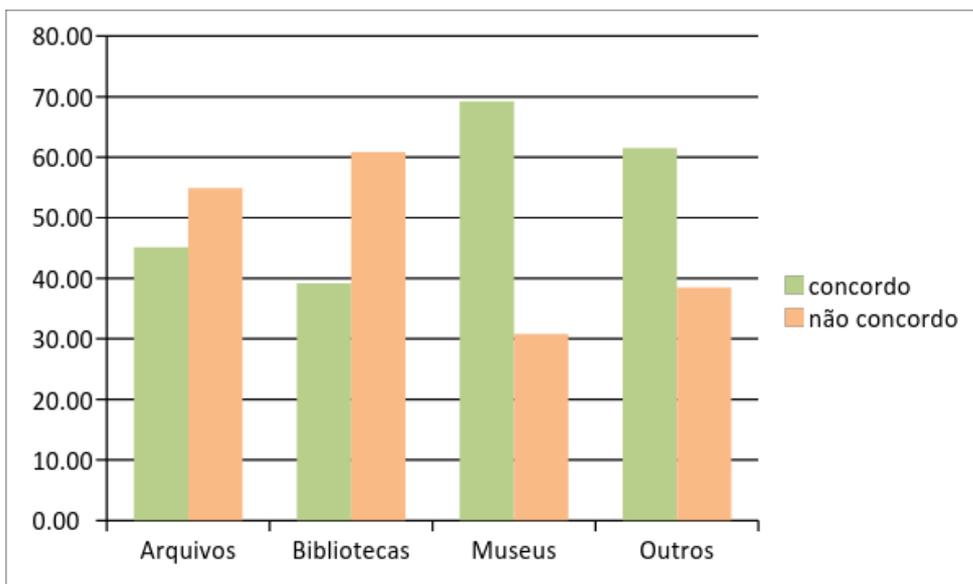


Gráfico 10 - Disponibilidade para confiar ODs a entidade terceira, pública ou privada, e digna de confiança

Conclui-se, pois, sobretudo tendo presentes as respostas obtidas na opção anterior, que para a maioria dos respondentes integrados nas CdP Arquivos e Bibliotecas, a confiabilidade da entidade terceira não se revela condição suficiente. Manifesta-se a preferência por uma entidade pública. O mesmo não se passa em relação aos respondentes integrados em Museus e ao grupo Outros, para a maioria dos quais esse aspeto é bastante menos valorizado.

No entanto os respondentes consideraram viável confiar os seus objetos digitais a uma entidade terceira desde que contratualizando um conjunto de garantias prévias.

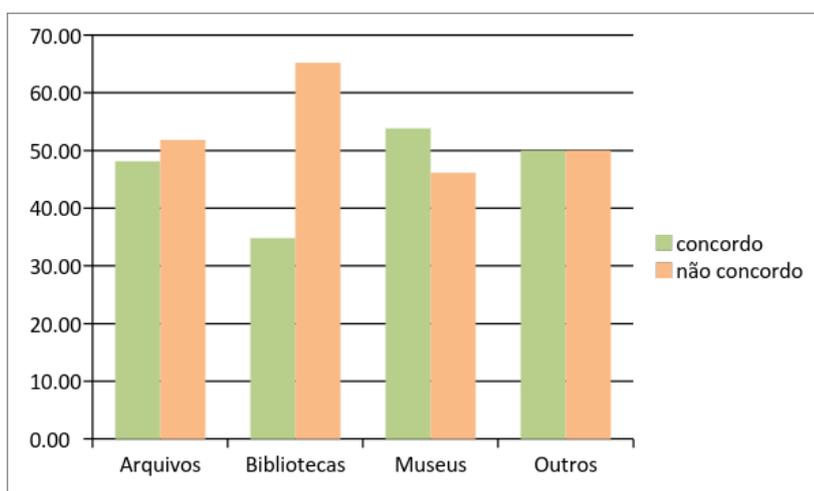


Gráfico 11 - Disponibilidade para confiar ODs a uma entidade terceira, desde que assegurado um conjunto mínimo de garantias a acordar previamente

De uma forma geral as CdP entendem que a gestão do património digital em comum tem vantagens financeiras, como se comprova pelas percentagens assumidas pelas opções de resposta “concordo” e “concordo inteiramente”, patentes no gráfico que se segue:

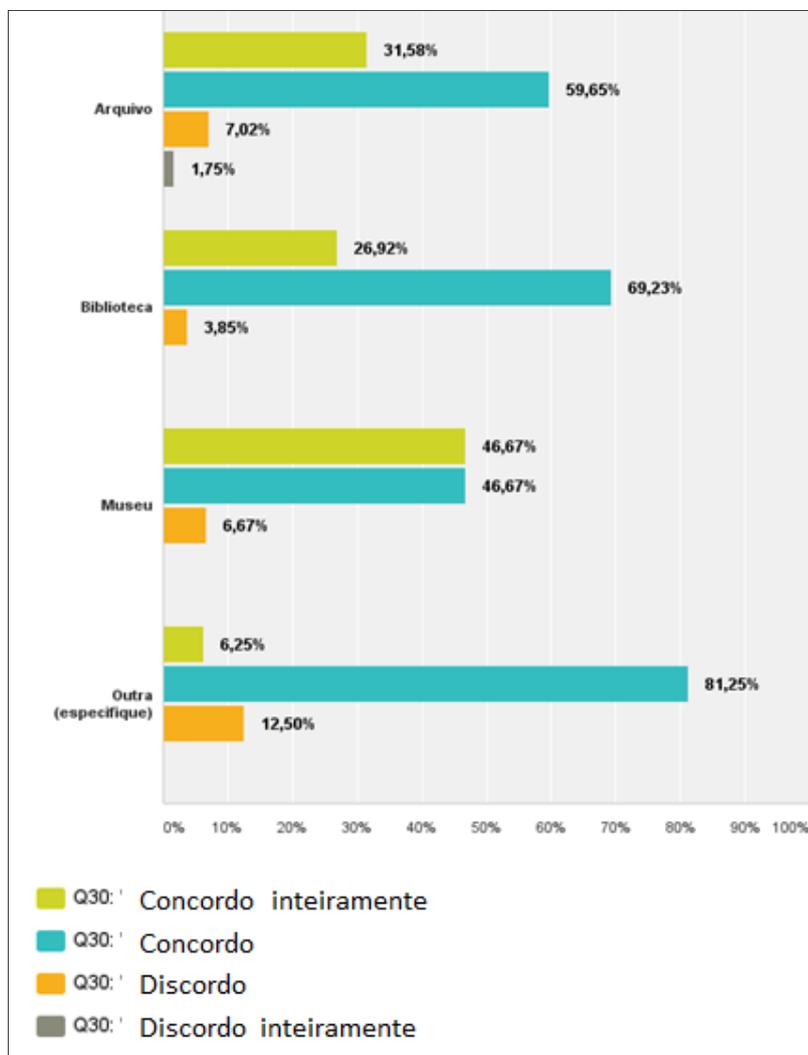


Gráfico 12 - Possibilidade construção de uma rede partilhada, tendo em vista a diminuição dos encargos financeiros

Existe disponibilidade por parte das instituições para assumir compromisso e responsabilidade relativamente a um modelo de gestão de tipo participativo.

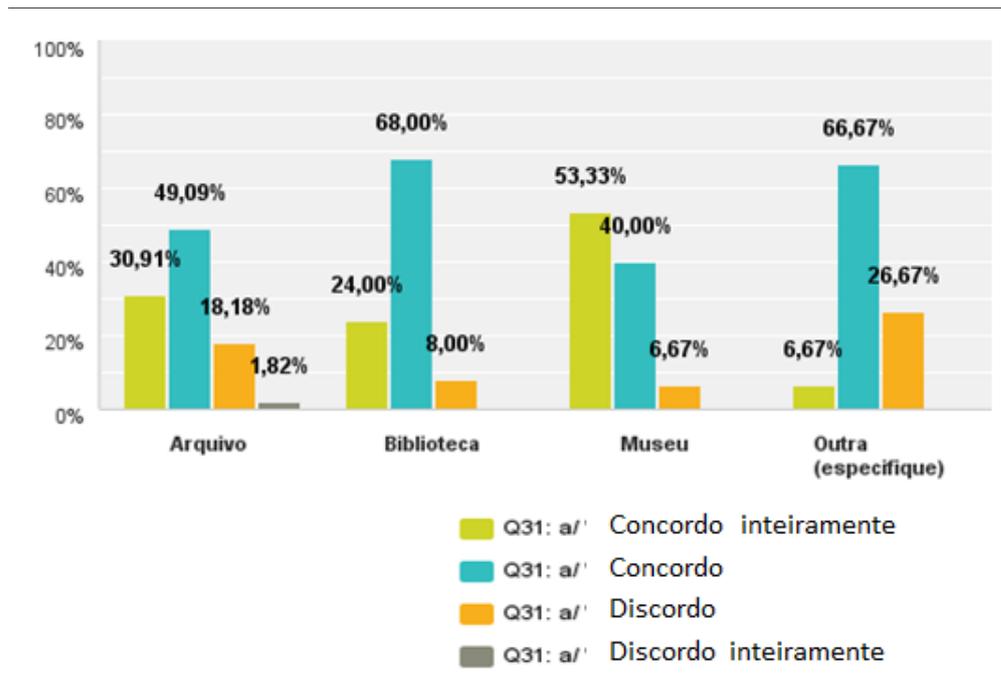


Gráfico 13 - Possibilidade de opção por um modelo descentralizado de gestão da rede

Relativamente ao modelo de exploração comercial proposto para assegurar a sustentabilidade financeira da RCPD - cedência de percentagem sobre a venda de cópias de objetos digitais - a maioria dos inquiridos das várias CdP concorda, sobretudo os das CdP Arquivos e Museus.

Constata-se que a maioria das CdP concorda com o conceito de um “repositório que receba e preserve diferentes tipos de património: de Arquivo, de Biblioteca, de Museu, etc.”, sendo irrelevante a percentagem dos respondentes que discordam inteiramente (CdP Arquivos) e pouco expressiva a percentagem de respondentes que discorda (com maior expressão nas CdP agrupadas em Outros, logo seguida dos Arquivos e Museus, e com menor expressão nas Bibliotecas).

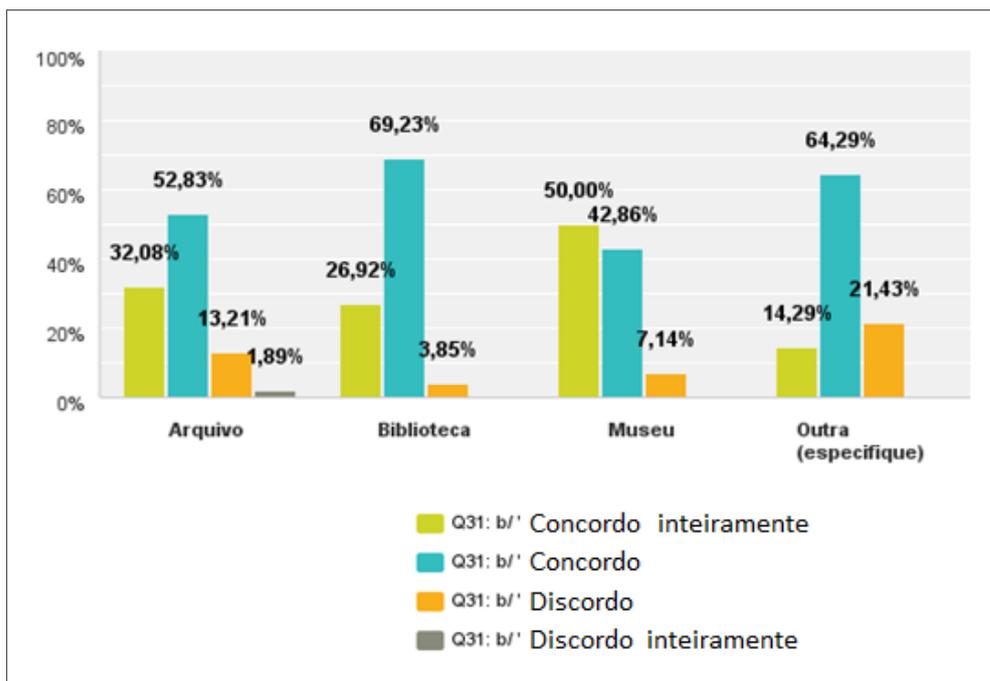


Gráfico 14 - Possibilidade de construção de uma rede transversal, aglutinando diferentes CdP

Apesar de algumas discordâncias - não muito relevantes -, conclui-se que existe uma vontade de preservar o digital, apostando numa conjugação de sinergias. O modelo proposto pode ser considerado, por isso mesmo, uma solução para contornar e ultrapassar os problemas de contexto da realidade portuguesa no que à área da preservação digital diz respeito, designadamente:

- A existência de um problema premente e transversal: urge preservar património digital;
- A ausência de mercado de tecnologias direcionadas à preservação digital.
- Os constrangimentos orçamentais das entidades públicas, mas também privadas, que dificultam uma abordagem individualizada do problema da preservação digital.

Para além destas conclusões o GT concluiu que:

Uma das questões em aberto passa por determinar da vantagem / necessidade de formalização, através de diploma legal, contendo, entre outros aspetos, a definição clara de uma entidade de gestão.

A clara interdependência entre o modelo de custos e o modelo de governação tem como consequência que os aspetos a aprofundar e a clarificar no que toca aos custos se repercutam necessariamente na proposta de governação.

As circunstâncias referidas poderão ser o melhor fundamento para justificar uma solução inovadora como a proposta.

Embora a proposta do modelo cooperativo, no entender do GT, tenha sido considerada globalmente bastante positiva e viável, existem, no entanto, aspetos que são suscetíveis de poder criar alguma **instabilidade** no funcionamento da rede:

- **O exercício da função executiva** não é facilmente regulável ou atribuível pelo que foram propostas hipóteses alternativas, no caso de não existir entidade voluntária.

- **Afetação de recursos humanos** à rede poderá ser resolvida por via da cedência de colaboradores das entidades aderentes ou por contratualização.
- No entanto, a situação administrativa pode ser problemática por pressupor questões burocráticas, sobretudo se as entidades aderentes forem públicas.
- **O exercício da função tecnológica** pode também não ser facilmente regulável, pressupondo proposta alternativa de contratualização de entidade externa, agravando custos.
- **A gestão financeira da rede** poderá obrigar à sua formalização legal ou algum instrumento legal que determine autorizações da tutela e/ou Ministério das Finanças.
- **O crescimento da rede** - n.º de aderentes superior a 12 - poderá exigir mecanismos de interação e de gestão mais complexos, eventualmente incompatível com um modelo de governação de inspiração cooperativa.

7 Perspetivas futuras

A elaboração de cenários futuros é o corolário do trabalho desenvolvido ao longo dos diversos passos do projeto. Pretende-se, em consonância com o plano inicial, definir cenários aceitáveis, tecnológica e socialmente, para o estabelecimento de uma rede comum para preservar património digital. Importa antes de mais analisar o nível de pontos convergentes que foi possível estabelecer ao longo dos trabalhos. Com efeito o ponto de partida assentava em 3 princípios unanimemente aceites pelos membros do GT (ver ponto 2 deste relatório). Ao longo do projeto foi possível constatar a existência de mais pontos de contacto, que no seu todo constituem o que consideramos ser uma forte base instalada para a constituição de uma rede cooperativa.

Por convergência não queremos dizer similitude. Pretendemos apenas exprimir a presença de aspetos comuns que possibilitam a interoperabilidade no que respeita a preservar informação digital. Da mesma forma os resultados sintetizados não expressam unanimidade mas sim maioria, muito significativa, na maior parte dos casos observados e contabilizados.

Neste sentido foi possível estabelecer pontos de convergência consolidados relativamente aos seguintes aspetos:

- Utilização de normas;
- Terminologias;
- Legislação;
- Valores partilhados no que respeita a propriedades expetáveis de ODs;
- Perceção comum no que respeita a relevância de preservar património digital;
- Disponibilidade para participar numa rede comum com informação multidomínio.

O diagrama seguinte representa o percurso crescente de convergência identificado, partindo de um ponto de partida comum e alargando-se progressivamente aos restantes pontos.

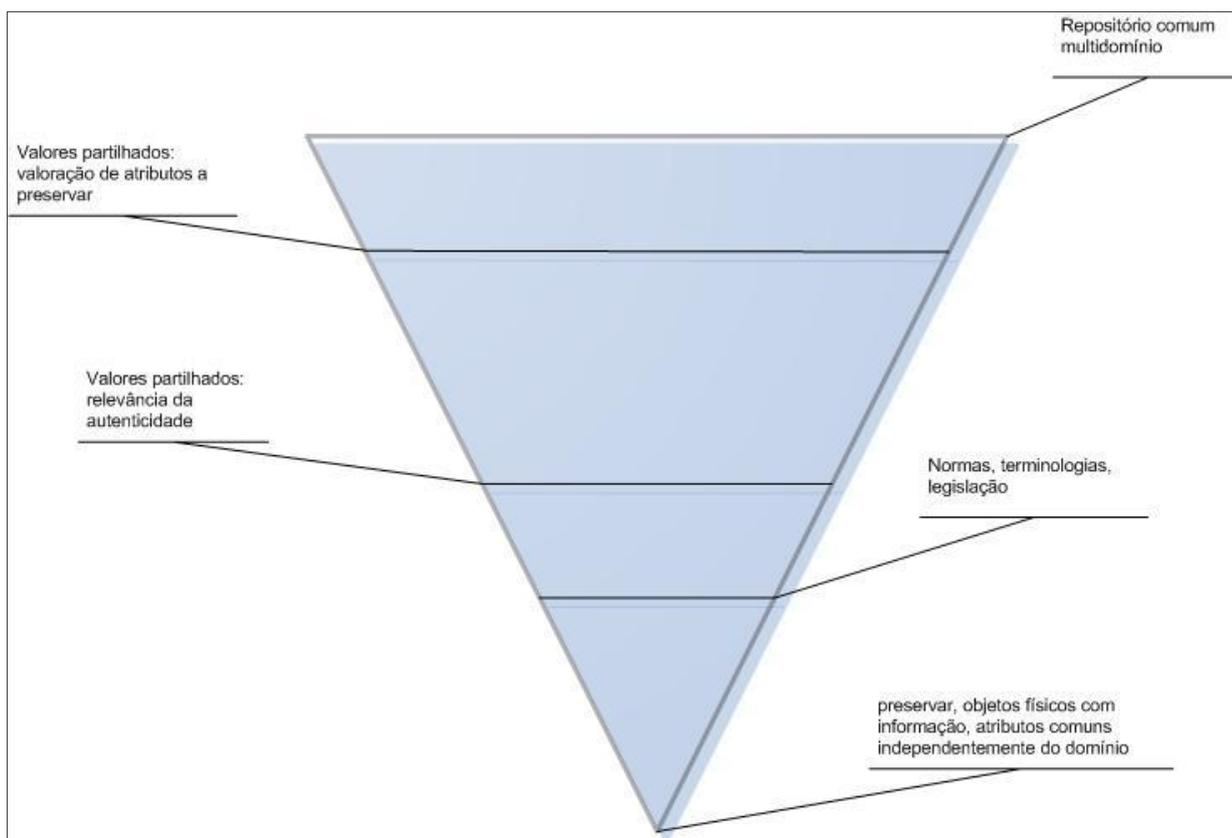


Diagrama 2 - Evolução de convergência

Não foi identificada convergência significativa nos seguintes pontos:

- Prática de avaliação, tópico onde se constatou uma situação muito dispersa e pouco consolidada;
- Disponibilidade financeira para sustentar uma rede de preservação com plataforma tecnológica comum.

Partindo do cruzamento e análise dos dados coletados é possível concluir ser necessário proceder a investigação adicional no que respeita a dois vetores.

1/ O primeiro é de natureza tecnológica. Um dos fatores que incrementa consideravelmente a estimativa de custos de um repositório comum é o armazenamento.

Quanto a este ponto foram identificadas situações extremas: há instituições, dentro de CdP, que possuem sistemas de armazenamento muito elaborados e de grande capacidade, normalmente proprietários. É, por exemplo, o caso das televisões. Simultaneamente há outros atores que não possuem sequer sistemas de armazenamento dedicados, recorrendo a sistemas de ficheiros em servidores ou a soluções tecnologicamente ainda menos elaboradas.

Esta situação leva-nos a considerar, na perspetiva de desenvolvimento de um repositório comum, a vantagem da coexistência de dois tipos de solução de armazenamento: comum e individualizado. Esta última hipótese implicaria o recurso à preservação digital distribuída. Isto significaria desenvolver um repositório centralizado contendo funcionalidades de preservação de acordo com o OAIS, mas que consiga aplicar remotamente essas funcionalidades, acedendo aos sistemas de armazenamento individualizados utilizados pelas instituições.

Esta abordagem, que significa na prática levar o *software* aos ODs, radica em experiências já consolidadas levadas a cabo nos EUA mas sobretudo no domínio de armazenamento e replicação de cópias de segurança de forma partilhada e distribuída. O que nos propomos realizar é um pouco mais elaborado, pois inclui o fornecimento de funções específicas de preservação digital que superam a função de armazenamento. Para atingir este cenário é necessária investigação. Embora a DGLAB esteja a desenvolver, no contexto de um projeto apoiado pelo QREN, a reestruturação do RODA de acordo com os resultados descritos neste relatório, este processo não contempla de forma totalmente integrada todas as vertentes desta nova filosofia de preservação digital. Trata-se, além disso, de uma área que implica a cooptação de outros atores e tecnologias, nomeadamente desenvolvedores de sistemas de armazenamento e *cloud*.

2/ O segundo é de natureza social/organizacional. Conclui-se ser necessário obter mais informação e conhecimento sobre as formas de articular organizações com culturas organizacionais e orçamentais díspares. A forma e conteúdo das relações a estabelecer entre atores é igualmente importante, particularmente considerando aspetos gestionários que podem ser simples numa organização singular, mas tornar-se complexos numa situação em que várias organizações pretendem cooperar e partilhar recursos financeiros.

Ainda neste domínio é igualmente imprescindível uma análise financeira e económica de forma a identificar formas possíveis de sustentação financeira da rede. Esta análise é importante por duas razões: 1/ o facto da maioria dos atores inquiridos não ter demonstrado disponibilidade para contribuir financeiramente para uma estrutura comum, o que determina procurar e encontrar alternativas que possibilitem uma inversão desta atitude, certamente tributária do presente estado de suborçamentação. 2/ tendo sido considerada como vantagem, ainda que parcial, uma independência financeira do Estado, recorrendo a fontes de receita alternativas, e ainda tendo em conta as experiências de estruturas similares, como os portais Europeia ou Apex, a investigação sobre este domínio do problema parece-nos essencial para a constituição de uma estrutura viável.

Considerando as necessidades e restrições acima descritas, o GT considerou que o cenário de continuidade mais adequado será a constituição de um consórcio com o objetivo de apresentar uma candidatura ao Horizonte 2020 para obter o financiamento necessário à prossecução destes propósitos. No entanto este cenário pressupõe a constituição de uma estrutura de suporte que atue como embrião de rede, prossequindo uma atividade em duas frentes:

1/ O suporte ao consórcio e a participação efetiva no projeto a constituir;

2/ A dinamização do público e de atores representantes de CdP de forma a:

- Disseminar de forma mais participativa os resultados do projeto;
- Promover conhecimento na área de preservação digital através de ações formativas e informativas;
- Obter uma base alargada para discussão e validação dos resultados do projeto;
- Cooptar massa crítica para a constituição efetiva da rede para preservação comum de património digital.

Considerando o exposto, o grupo executivo delineou uma estrutura possível de rede para suportar os objetivos acima descritos. O diagrama seguinte constitui a sua denotação gráfica:

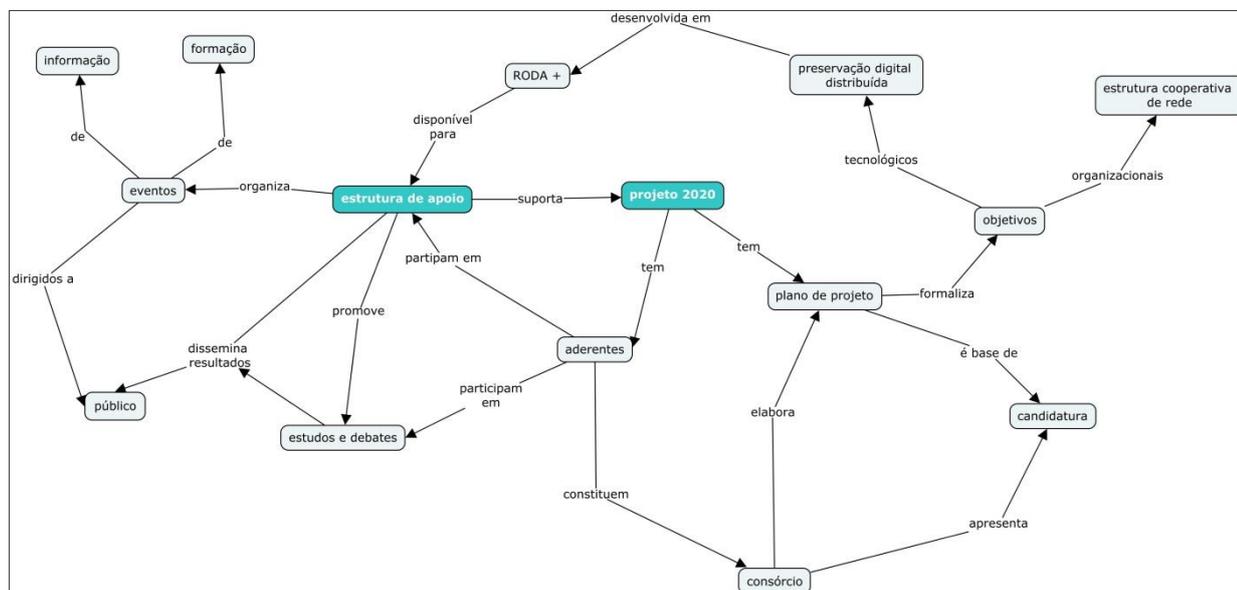


Diagrama 3 - Eventual estrutura a adotar pela rede

A matriz explicativa de cada uma das classes e relações é seguidamente apresentada:

Classes	Descrição	Relações
Estrutura de apoio	Conjunto de pessoas e organizações que apoiam a rede constituída na sua vertente de candidatura 2020 e na de disseminação de resultados e dinamização de conhecimento	Projeto 2020; Eventos; Público; Estudos e debates
Projeto 2020	Um projeto aprovado pelo 2020 constitui a base de construção e desenvolvimento do conhecimento e das ferramentas necessárias para construção de uma rede cooperativa	Estrutura de apoio; Aderentes; Plano de projeto

	para preservação digital	
Plano de projeto	Plano formalizado do projeto a entregar aquando da submissão de candidatura	Objetivos; Projeto 2020; Consórcio
Consórcio	Conjunto de organizações que configuram aderentes, tanto ao projeto 2020 como à estrutura de apoio. A participação nas duas vertentes não é vinculativa, podendo haver aderentes que participam numa delas ou em ambas simultaneamente. O consórcio deverá ter obrigatoriamente aderentes internacionais.	Aderentes; Plano de projeto; Candidatura
Candidatura	Ação final que culmina o processo de elaboração de formação de consórcio e elaboração do plano de projeto	Consórcio; Plano de projeto
Objetivos	Domínios de investigação do projeto	Plano de projeto; Estrutura cooperativa de rede; Preservação digital distribuída
Estrutura cooperativa de rede	Um dos domínios de investigação do projeto. Pressupõe a vantagem de uma estrutura social cooperativa compreendendo estudos e soluções abrangentes mas detalhadas quanto aos modelos de governação e de sustentação financeira mais adequadas numa perspetiva de grande diversidade organizacional e jurídica	Objetivos
Preservação digital distribuída	Um dos domínios de investigação do projeto. Pressupõe a vantagem de levar a computação que suporta funções normalizadas de preservação digital aos repositórios de armazenamento onde residem os objetos a preservar. Exige articulação com uma rede ubíqua e heterogénea de sistemas de armazenamento.	Objetivos; RODA +
RODA +	O arquivo digital português desenvolvido pela DGLAB com base no OAIS. Neste momento está a ser reestruturado ao abrigo de um projeto QREN. Permite a filosofia de preservação digital distribuída e preservação digital multidomínio. Pode ser usado e testado livremente pela estrutura de apoio.	Preservação digital distribuída; Estrutura de apoio
Aderentes	Pessoas coletivas e individuais que queiram participar no consórcio e/ou na estrutura de apoio	Estrutura de apoio; Consórcio; Projeto 2020

Eventos	Ações de dinamização e promoção de conhecimento organizadas pela Estrutura de apoio. Destinadas ainda à disseminação maciça dos resultados do projeto 2020. Estes eventos destinam-se ao público em geral, embora possa vir a ser efetuada segmentação.	Estrutura de apoio; Informação; Formação; Público
Formação	Um tipo de evento destinado a ministrar formação em áreas relacionadas com preservação digital	Eventos
Informação	Um tipo de evento em que é fundamentalmente trocada e discutida informação. São normalmente workshops, congressos, seminários	Eventos
Público	Toda e qualquer pessoa, profissional ou não, interessada nas temática de preservação digital e património	Estrutura de apoio; Eventos
Estudos e debates	Programas destinados a aderentes consistindo em temáticas específicas discutidas de forma orientada e contextualizando práticas estabelecidas de comunidades profissionais	Estrutura de apoio; Eventos

Quadro 15 - Matriz explicativa das classes e relações utilizadas

8 Anexos

Considerando a dimensão dos anexos, preferiu-se indicar o link para o local onde podem ser consultados e descarregados.

8.1. Legislação

Consultar anexo em WWW:<URL: http://1seminariopreservacaopatrimonioidigital.dglab.gov.pt/wp-content/uploads/sites/19/2015/11/1_Anexos_Quadro_Reg_final.xlsx>.

8.2 Terminologia

Consultar anexo em WWW: <http://1seminariopreservacaopatrimonioidigital.dglab.gov.pt/wp-content/uploads/sites/19/2015/11/1_Anexos_Quadro_Reg_final.xlsx>.

8.3. Normas e mapeamentos de normas

Consultar anexo em WWW:<URL: http://1seminariopreservacaopatrimonioidigital.dglab.gov.pt/wp-content/uploads/sites/19/2015/11/1_Anexos_Quadro_Reg_final.xlsx>.

8.4 Formatos

Consultar anexo em WWW:<URL: http://1seminariopreservacaopatrimonioidigital.dglab.gov.pt/wp-content/uploads/sites/19/2015/11/1_Anexos_Quadro_Reg_final.xlsx>.

8.5 Folha de cálculo de custos de serviços de preservação digital

Disponível em WWW:<URL: http://1seminariopreservacaopatrimonioidigital.dglab.gov.pt/wp-content/uploads/sites/19/2015/10/CustoPD_NumeroOD.xlsx >.

8.6 Informação adicional

Neste anexo são referidas todas as iniciativas empreendidas ao longo do projeto. Estas encontram-se também no respetivo sítio web. A sua compilação neste anexo tem como único propósito facilitar a sua consulta.

Eventos de disseminação organizados pelo GT (SOS digital)

#	Data	Tema
11	18/06/2015	Continuidade digital: Património sonoro e fonográfico
10	20/05/2015	Sessão integrada nas jornadas “Arquivos Memória e Futuro” organizadas pelo Arquivo distrital de Aveiro
9	27/04/2015	Armazenamento de dados numa perspetiva de preservação digital a longo prazo
8	15/01/2015	Património musical e preservação digital
7	9/12/2014	Continuidade digital: Atualização do projeto
6	10/07/2014	Gestão de dados de investigação
5	26/06/2014	Continuidade digital (especial): Problemas técnicos específicos da preservação digital
4	18/06/2014	Jornada: Debater, preservar património digital (apresentar, esclarecer e debater conceitos base e avançados sobre vários aspetos e perspetivas de preservação digital com o público interessado)
3	29/04/2014	Continuidade digital: Análise comparativa de normas de metadados e apresentação preliminar do inquérito sobre perceção de autenticidade de informação digital
2	31/03/2014	Continuidade digital: Legislação, terminologia, normas, formatos e esquemas de metadados
1	26/02/2014	Continuidade digital: Apresentação do projeto

Comunicações apresentadas em eventos

BARBEDO, F.; RODRIGUES, A.; RUNA, L.; SANT’ANA, M. - *Preserving digital heritage: a network centric approach*. DLM-Forum 2014.

BARBEDO, F.; RODRIGUES, A.; RUNA, L.; SANT’ANA, M. - *Preservar património digital através de redes de cooperação multidomínio*. Encontro Arquivos Científicos, 2014-07-04 - Lisboa, FCSH-UNL - IHC-FCSH & ACT-FCT.

BARBEDO, F.; RODRIGUES, A.; RUNA, L.; SANT’ANA, M. - *Redes, Galáxias, Património Digital: Importa-se de repetir?* Comunicação aceite para o Congresso BAD 2015.

BARBEDO, F.; RODRIGUES, A.; RUNA, L.; SANT’ANA, M. - *Rede de Preservação Comum do Património Digital: custos, proveitos e balanços*. Comunicação aceite para o Congresso BAD 2015.

Documentos de projeto

#	Título	Resumo
1	Relatório de projeto - passo 1 e anexos.	Neste relatório foram comparadas e analisadas legislação, terminologias, normas de representação e formatos reportados pelos membros do Grupo de Trabalho. Apresenta a metodologia de análise seguida para cada atividade bem como as conclusões obtidas e os anexos.
2	Relatório estatístico sobre percepção de autenticidade e avaliação de objetos digitais.	Este relatório, desenvolvido no âmbito do passo 2, recolheu as opiniões do grupo de trabalho sobre os aspetos acima referidos. Importa perceber em que medida e grau a autenticidade é um valor partilhado pelas comunidades de prática representadas Pretendia-se também apurar se as diferentes comunidades de prática avaliam objetos digitais antes destes serem integrados no universo patrimonial, bem como perceber que critérios são usados.
3	Simulador de custos de preservação digital. (excel).	Esta ferramenta foi desenvolvida no âmbito do projeto e permite simular os custos que uma determinada instituição teria para reservar a sua informação digital. Conforme se refere no relatório final, é necessária investigação adicional para estimar custos de forma assertiva, pois estes estão dependentes de outras variáveis que carecem de ser equacionados em situação real. O cálculo é efetuado com base num cenário elaborado pelo GT o qual pode ser alterado com impacto no custo final. Mesmo assim a presente ferramenta permite obter uma estimativa da grandeza desses custos.
4	Conclusões preliminares do Inquérito Público - Passo 6.	Este inquérito pretendeu obter informação sobre princípios e métodos utilizados para integrar recursos digitais em universos patrimoniais e apurar a viabilidade da criação de um repositório comum de preservação de património digital.
5	Relatório estatístico final do inquérito público - Passo 6.	Este inquérito pretendeu : obter informação sobre princípios e métodos utilizados para integrar recursos digitais em universos patrimoniais; apurar a viabilidade da criação de um repositório comum de preservação de património digital
6	Relatório final do projeto.	Balanço de todas as atividades desenvolvidas no projeto e respetivos resultados e conclusões. Propõem-se cenários de continuidade e de investigação adicional
7	Percepção da Autenticidade de Sistemas de Informação de Entidades da Administração Pública Central	Neste inquérito pretende-se medir e avaliar o grau de percepção de diversos tipos de utilizadores de Sistemas de Informação de organismos da Administração Pública Central relativamente à autenticidade da informação digital com que trabalha quotidianamente, por forma a inferir tendências de opinião sobre a realidade em causa

Bibliografia consultada

#	Descrição
1	Australia. National Library - <i>Guidelines for the preservation of digital heritage</i> . (CI-2003/WS/3). UNESCO, 2003.
2	Ball, Alex - <i>Preservation and Curation in Institutional Repositories</i> . Version 1.3. Edinburgh, UK: Digital Curation Centre, 2010.
3	Blue Ribbon Task Force on Sustainable Digital Preservation and Access - <i>Sustaining the digital investment: Issues and challenges of economically sustainable digital preservation: Interim report</i> . 2008.
4	Briggs, Mark - <i>Journalism 2.0. How to Survive and Thrive: A digital literacy guide for the information age</i> . 2007.
5	Commission of the European Communities - <i>Commission Staff Working Document on the implementation of the Recommendation of the European Parliament and Council of 16 November 2005 on film heritage and the competitiveness of related industrial activities</i> . Brussels: Commission of the European Communities, 2008.
6	Di Carlo, Annarita; Picciotti, Gianluca; Bartoleschi, Monica; Lucaferri, Lara - <i>Glossary of Rights</i> . Presto Centre, 2011.
7	Digital Preservation Coalition - <i>Preservation Metadata</i> . 2nd edition. Digital Preservation Coalition. ISSN 2048-7916. (2013).
8	Digital Preservation Europe - <i>UMID - Unique Material Identifier: briefing paper</i> .
9	European Broadcast Union - <i>International Broadcast Tape Number (IBTN). Bar-code labels and register of Facilities Codes and Media Carriers: Technical Specification</i> . (N.º EBU - TECH 3279). Geneva: EBU, 2011.
10	European Broadcast Union. European Broadcast Union - <i>EBU Core Metadata Set (EBUCore)</i> . Version 1.4. (N.º.EBU - TECH 3293). Geneva: EBU, 2013.
11	European Broadcast Union. Operating Eurovision and Euroradio - <i>P_Meta. Metadata Library. Specification 2.2</i> . (No. EBU - TECH 3295). Geneva: EBU, 2011.
12	European Research Consortium for Informatics and Mathematics - <i>ERCIM News. Special theme: Digital Preservation</i> . ERCIM EEIG. ISSN 0926-4981. 80 (2010).
13	International Federation of Television Archives - <i>Recommended standards and procedures for selection and preservation of television programme material</i> . IFTA, 1996.
14	International Standard Audiovisual Number - <i>ISAN Metadata Schema</i> . Version 1.12. Geneva: ISAN, 2012.
15	Kyong-Ho Lee; Slattery, Oliver; Richang Lu; Xiao Tang; McCrary, Victor - "The State of the Art and Practice in Digital Preservation". In <i>Journal of Research of the National Institute of Standards and Technology</i> . 107, 1 (January-February 2002).

16	Millar, Laura - <i>Authenticity of Electronic Records: A Report Prepared for UNESCO and the International Council on Archives (ICA Study 13-2)</i> . ICA, 2004.
17	Online Computer Library Center. Center for Research Libraries - <i>Trustworthy Repositories Audit & Certification: Criteria and Checklist. Version 1.0</i> . 2007.
18	Peacefulfish - <i>Digital agenda for the European film heritage: Challenges of the Digital Era for Film Heritage Institutions: Final report</i> . Study prepared for the European Commission, DG Information Society and Media. Luxembourg: Publications Office of the European Union, 2012. ISBN 978-92-79-22809-4.
19	PRESTO - Preservation Technologies for European Broadcast Archives - <i>D7.2 Guide to Metadata in Preservation</i> . (PRESTO-W7.2-BBC-020430 Metadata Reference Process). 2002.
20	Riley, Jenn - <i>Glossary of Metadata Standards</i> . Jenn Riley, 2009-2010 (cop.).
21	Riley, Jenn - <i>Seeing Standards: A Visualization of the Metadata Universe</i> . Jenn Riley, 2009-2010 (cop.).
22	Skinner, Katherine; Schultz, Matt (Ed.) - <i>A Guide to Distributed Digital Preservation</i> . Atlanta: Educopia Institute, 2010 (cop.). ISBN: 978-0-9826653-0-5.
23	Torres, Ana Maria Henrique Franqueira - <i>O Sistema de Gestão e Arquivo de Conteúdos da SIC - Sociedade Independente de Comunicação, S.A.: proposta de indicadores para medir a eficiência de um arquivo digital audiovisual, com base na análise de valor</i> . Tesis Doctoral. Universidad de Alcalá, Departamento de Filología, Documentación y Comunicación Audiovisual. Alcalá de Henares, 2014.
24	United Kingdom. National Archives - <i>Digital Preservation Policies: Guidance for archives</i> . United Kingdom: National Archives, 2011.
25	United Nations Educational Scientific and Cultural Organization - <i>The Memory of the World in the Digital Age: Digitization and Preservation</i> . An international conference on permanent access to digital documentary heritage. Luciana Duranti e Elizabeth Shaffer (Eds.). UNESCO, 2013.
26	University of California Libraries. Next-Generation Technical Services - <i>Digital Asset Management System (DAMS) Requirements: final report</i> . California: University of California July 20, 2012.
27	Lei n.º 107/2001 . In Diário da República n.º 209/2001, Série I-A de 2001-09-08. Estabelece as bases da política e do regime de proteção e valorização do património cultural.