

Preservação Digital

Vítor Lopes¹

1) Universidade do Minho, Guimarães, Portugal

vtr_lps@hotmail.com

Resumo

Cada vez mais, e com o passar do tempo existe a necessidade de se preservar a informação que criamos no dia-a-dia. No presente, a criação de informação tem crescido de forma exponencial, o que nos leva a pensar em meios para a preservar de modo a mantê-la ao longo do tempo. Com isso, a preservação digital tem vindo a ganhar relevo ao longo do tempo, sendo que no futuro se tornará um factor indispensável para humanidade. Este artigo tem como principal objectivo, dar uma visão global do que é a preservação digital, de como se aplica e quais os requisitos para a sua implementação. Será também explicado o que é a preservação, seguido de termos indispensáveis para a mesma, como é o caso dos repositórios institucionais, o modelo de referência *Open Archival Information System (OAIS)*, os objectos digitais (nato-digitais e digitalizados) bem como os seus formatos e suportes. Por último este artigo faz um enquadramento da preservação digital no seio das organizações, com o objectivo de mostrar a importância que esta pode ter no desempenho futuro das mesmas.

Palavras-chave: preservação, formatos, suportes, informação, tecnologia, organizações, estratégias, repositórios, objectos digitais.

1. Introdução

A preservação digital aparece hoje em dia como a solução em termos de preservação de informação para o futuro.

A preservação sempre teve particular importância no mundo desde a criação de informação da idade média até aos dias de hoje. Sem a preservação, hoje nunca poderíamos ter acesso a documentos, livros, gravuras entre outra informação que vem sendo transmitida de geração em geração. Esta problemática da preservação existente também nos dias de hoje no que diz respeito à preservação digital, futuramente continuará a existir com um outro tipo de preservação adequado à sua era [Hedstrom e Montgomery 1998].

Tal como aconteceu com os nossos antepassados, a nossa preocupação hoje em dia é a de conseguir garantir que a informação que temos estará disponível tanto em termos de acesso como de integridade ao longo dos tempos. Para isso contamos com estudos e metodologias que têm sido feitas e que continuarão a ser feitas com o propósito de sempre, e também optamos por uma condução do processo de implementação.

Hoje em dia estamos particularmente atentos às novas tecnologias e à forma como devemos dar utilidade de modo a conseguirmos retirar todas as suas potencialidades no contexto da preservação da informação. Daí ter surgido o nome digital que visa a era tecnológica corrente e a possibilidade de se conseguir manter a informação em formato digital tirando partido das tecnologias de informação.

O problema que se coloca tem a ver com os formatos, tamanhos e dispositivos de leitura que sejam capazes de serem utilizados com um tempo de vida longo, sem se tornarem obsoletos e que consigam manter a informação íntegra de maneira que o receptor consiga interpretar a mensagem que o emissor transmite [Thomaz 2004].

2. O que é a preservação?

A maior parte dos colecionadores sabem que a organização e armazenamento físico de materiais é susceptível de proporcionar uma maior garantia de recuperação do que simplesmente atirar esses mesmos materiais para cima de uma mesa. Isto é o equivalente a depositar um artigo electrónico num repositório institucional (*IR*), ao contrário, de simplesmente fazer o *uploading* desse artigo para um servidor ou página Web pessoal.

A organização do *IR*, juntamente com o empenho da instituição patrocinadora irá proporcionar uma maior garantia do que um servidor com um suporte desconhecido.

O simples armazenamento dificilmente é uma solução completa, no entanto, a longo prazo, até mesmo materiais físicos podem sofrer alguma degradação física. O equivalente processo de degradação de materiais digitais, geralmente, é causado pela obsolescência dos formatos, devido a mudanças na tecnologia de aplicações de *software*. Muitas vezes esta mudança é considerada um processo rápido em comparação com a degradação física dos materiais, e podem ser melhorados por processos técnicos especializados, tais como a migração dos formatos. Em ambos os casos, os riscos para a degradação de materiais físicos e digitais podem ter uma melhor relação custo/eficácia, sendo abordado por um especialista de preservação [Maniatis et al 2005]. Para materiais informativos como livros e jornais, esses serviços foram tradicionalmente oferecidos por bibliotecas e arquivos, e tais organizações podem ser bem posicionadas para servir de fontes digitais [Hitchcock et al 2007].

A preservação digital abrange assim uma vasta gama de actividades, desde o armazenamento até à transformação, dependendo da natureza dos recursos e da fonte, tal como a gama de serviços de preservação pode ser igualmente ampla.

Como tais serviços não são ainda muito praticados ou avaliados, um ponto de partida útil é considerar o que é conhecido acerca do conteúdo alvo do *IR*.

3. Repositórios Institucionais (IR's)

Um Repositório Institucional é um local online para recolha, preservação e difusão de informação em formato digital. Normalmente, o resultado intelectual de uma instituição é, nomeadamente um centro de investigação.

Esta é uma definição simplista de ver o que é um *IR*, dado que hoje em dia existem variadíssimas implementações de *IR's*, pertencendo a maioria delas a instituições universitárias. O impulso para a criação de *IR's* foi dado pela *Open Archives Initiative (OAI)*, em 1999.

A *OAI* é um movimento que desenvolve e promove padrões de interoperabilidade entre repositórios digitais. Assim, ela criou o protocolo *OAI-PMH*, um meio simples de se compartilharem metadados entre servidores distribuídos [Ferreira 2007].

O movimento *OAI* tem como principal objectivo, estender o acesso a repositórios de artigos científicos de modo a aumentar a disponibilização desses mesmos artigos. Embora o principal objectivo da *OAI* sejam os repositórios científicos, ela tem também em atenção os repositórios de músicas, *URL's* e vídeos.

Resumindo, podemos dizer que o papel da *OAI* visa a criação de protocolos que proporcionem a construção de repositórios de informação digital.

Para termos uma ideia do que se tem vindo a fazer em termos de *IR's* podemos sempre consultar o *Registry of Open Access Repositories (ROAR)* [*Registry of Open Access Repositories (ROAR)* 12/05/2008] onde estão registados os variados *IR's* existentes. No *ROAR* podemos ter acesso a estatísticas sobre o funcionamento dos repositórios, bem como das suas políticas através do *Registry of Open Access Repository Material Archiving Policies (ROARMAP)* [*Registry of Open Access Repository Material Archiving Policies (ROARMAP)* 12/052008].

Na figura 1, podemos ver a disposição global do número de repositórios espalhados pelo mundo.

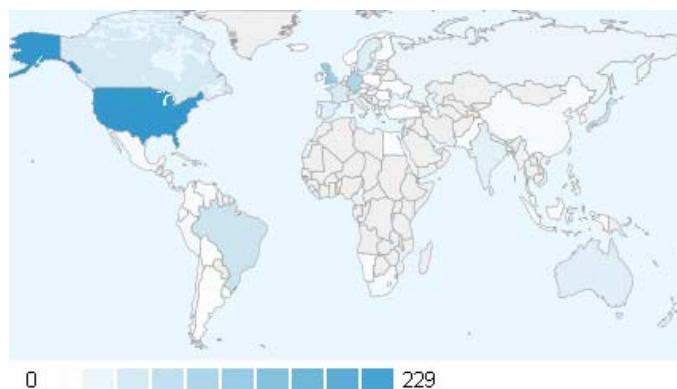


Figura 1 – Número de repositórios no Mundo

Para além do ROAR temos também o *Open Directory of Open Access Repositories (OpenDOAR)* [*Open Directory of Open Access Repositories (OpenDOAR)* 12/05/2008], onde também temos acesso às mais variadas estatísticas sobre os repositórios existentes em todo o mundo. Por exemplo na figura 2 podemos consultar os tipos de conteúdo nos repositórios *OpenDOAR* no mundo inteiro.

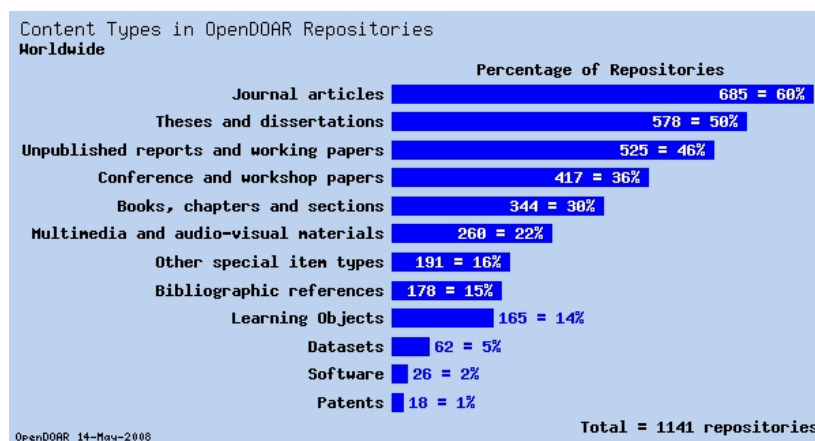


Figura 2 - Tipos de conteúdo nos repositórios OpenDOAR a nível Mundial

Para além da necessidade de seguir os protocolos da *OAI* na construção de um repositório, temos que ter também em atenção qual ou quais os *softwares* que vão lidar com o repositório e consequente informação, tais como:

- *DSpace*
- *Eprints*
- *Bepress*

4. Modelo de referência *Open Archival Information System (OAIS)*

Um *OAIS* é um arquivo, constituído por uma organização de pessoas e de sistemas, que aceitou a responsabilidade de preservar informação e torná-la disponível para uma Comunidade Designada [CSSDS 2002].

Este modelo foi criado pela *National Aeronautics and Space Administration (NASA)* através do seu *Consultative Committee for Space Data Systems (CCSDS)*, sendo ele também uma norma *International Organization for Standardization (ISO)* correspondente ao número 14721:2002. Esta descreve um enquadramento conceptual para um repositório de dados genérico, aberto a todas as comunidades com garantias de confiança [Baptista et al. 2007]. O modelo *OAIS* pode ser aplicado a qualquer arquivo, embora seja especificamente mais aplicável às organizações com a responsabilidade de tornar a informação disponível a longo prazo. No entanto é

importante salientar que deverá existir em simultâneo com o modelo *OAIS*, um modelo de informação onde se encontram descritos os requisitos de metadados de preservação a longo prazo [Saramago 2004].

O modelo *OAIS* pode ser visionado na figura 3 que é apresentada em seguida [Thomaz e Soares 2004].

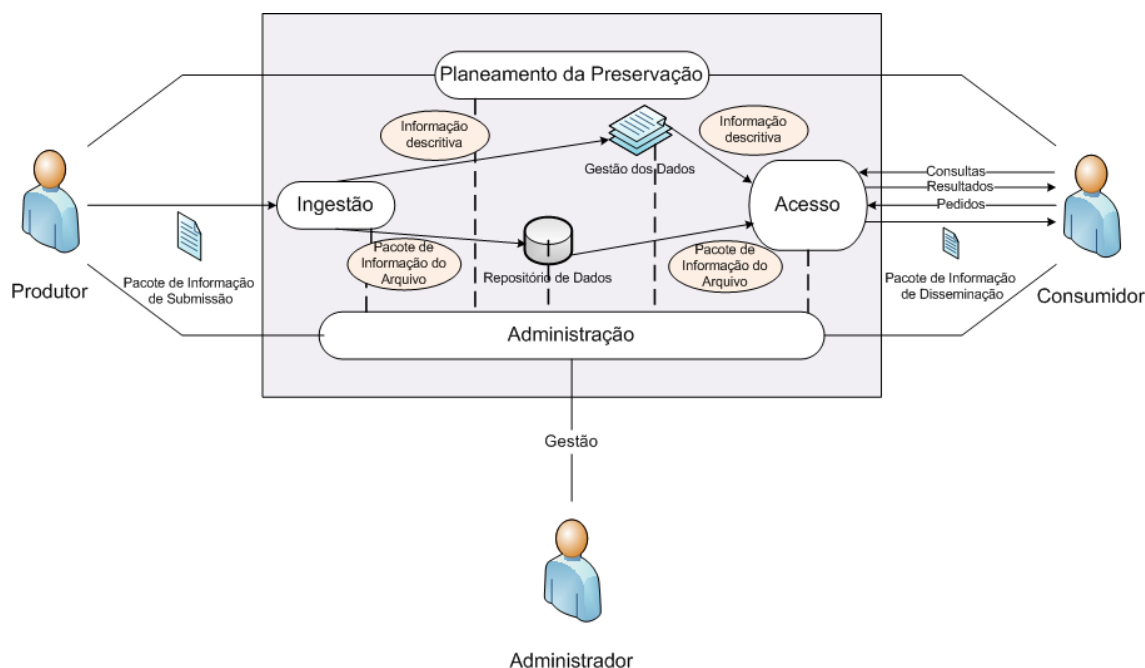


Figura 3 – Modelo de Referência *OAIS*

O produtor, que pode ser ou não o criador da informação assume o papel de fornecedor da informação a ser submetida no modelo. O administrador ou gestor do modelo tem o papel de impor a política *OAIS*, tendo a responsabilidade de monitorizá-la passo a passo de modo a ser cumprida. Por último temos o papel do consumidor ou utilizador, que é aquele que interage com o *OAIS* de forma a obter a informação que procura. O modelo *OAIS* só funciona na presença destas três personagens essenciais.

O processo em torno do *OAIS* desenrola-se da seguinte forma: o produtor insere o pacote de informação de submissão na ingestão onde através do controlo da administração irá dividir esse pacote em duas partes. Uma parte é o pacote de informação do arquivo que será inserido no repositório de dados, a outra parte é a informação descritiva que irá para a gestão dos dados, tendo em conta o planeamento da preservação e respectivos requisitos dos metadados. Para o acesso por parte do consumidor/utilizador ao pacote de informação de disseminação, é necessário proceder-se à junção do pacote da informação do arquivo com a sua referente

informação descritiva, de maneira a ser criado o pacote de informação de disseminação para posteriormente ser entregue ao consumidor/utilizador.

5. Objectos Digitais

Um documento electrónico ou documento digital, é todo o registo gerado ou recebido por uma entidade pública ou privada, no desempenho das suas actividades, podendo ser armazenado e disponibilizado ou não, através de sistemas de computação [Proença e Lopes 2004]. Esta descrição do que é um objecto digital também pode ser vista de outra forma, em que um objecto digital pode ser definido como todo e qualquer objecto de informação que possa ser representado através de uma sequência de dígitos binários [Thibodeau 2002].

No entanto, os objectos digitais estão divididos em dois tipos distintos:

- Nado-digitais
- Digitalizados

Os objectos digitais do tipo nado-digitais são objectos que já são criados em suporte digital, como por exemplo, um documento em Word ou em Excel, uma fotografia digital, ou qualquer outro tipo de documento que seja digital à nascença.

Os objectos digitais do tipo digitalizados são objectos que provêm de objectos físicos, ou seja, são digitalizados a partir de objectos físicos, como por exemplo, a digitalização através de um scanner de um documento que se encontre em papel. Outro modo de tornar esse mesmo objecto físico em objecto digital, pode ser por exemplo, tirar uma fotografia através de uma máquina digital ao objecto físico.

No entanto, temos de realçar o facto de que é necessário manter a integridade do objecto físico no objecto digital, isto de forma a este poder ser preservado e posteriormente observado, garantindo a sua perfeita interpretação.

Os documentos digitalizados, subdividem-se em três classes de informação:

- Gráfica
- Textual
- Sonora

A informação gráfica diz respeito a imagens digitalizadas, podendo estas serem obtidas através de vários mecanismos, como por exemplo, um scanner, uma máquina fotográfica digital ou também por uma máquina de filmar digital. Estes objectos podem ser guardados tanto em estado

digital num computador, como em formato físico (isto depois de ter sido imprimido através de uma impressora).

Os textos digitalizados podem ser criados através da digitalização de texto, que pode ser feito de duas maneiras diferentes: semi-automática ou automática, através do *Optical Character Recognition (OCR)* que identifica texto em imagens digitais. Este tipo permite o processamento automático dos conteúdos, como por exemplo, a tradução automática, a pesquisa rápida, a indexação, etc.

Os documentos digitalizados sobre a classe sonora, são obtidos através de textos lidos, declamados, cantados, de músicas, etc. Isto consiste na gravação directa de um formato digital de som ao vivo, ou pela transmissão para formato digital de som já existente por outras técnicas, como por exemplo, analógicas, meios mecânicos ou electromagnéticos [Proença e Lopes 2004].

Contudo, um documento digital só é útil, desde que seja fiel ao original, para que a mensagem que é percebida através do contacto com ele, seja igual à da interpretação do objecto original.

6. Formatos e suportes

Não faria sentido falar de objectos digitais sem me pronunciar sobre os seus formatos e suportes. Esta é uma problemática de grande importância para a preservação digital, pois tanto os formatos como os suportes para os objectos digitais estão em constante evolução, fazendo com que rapidamente os formatos e suportes se tornem obsoletos, acabando por se tornarem um grande entrave à preservação digital [Thomaz 2004].

Contudo, a preservação digital tem de ser capaz de responder positivamente a essas alterações fruto da evolução, de modo a garantir que os objectos anteriormente criados em outros formatos e suportes sejam lidos e decifrados, fazendo com que a sua mensagem chegue na perfeição ao consumidor final. Para isso foram criadas e ainda continuam a serem criadas estratégias de preservação para fazer com que apesar dos formatos, suportes e tecnologia evoluírem e se alterarem, os objectos tenham capacidades e meios de serem lidos e que as suas mensagens cheguem ao consumidor final.

Falando mais especificamente dos formatos, estes têm vindo a mudar ao longo do tempo como consequência da evolução tecnológica, como podemos observar na evolução dos ficheiros de imagens que podem ser encontrados em variadíssimos formatos diferentes e distintos, como por exemplo, *JPEG*, *PNG*, *GIF*, etc.

Ainda mais significativa é a mudança que tem vindo a ser registada nos formatos de suporte físico, tendo este evoluído muito em pouco tempo, como podemos observar na evolução desde as disquetes para os *CD's*; dos *CD's* para os *DVD's*; mais recentemente com a evolução dos *DVD's* para as memórias *flash* com principal atenção para as *pendrives (USB Flashdisk)* e também para os cartões de memória, como podemos ver na figura 4.

Apesar destes já existem há já algum tempo, têm agora ganho mais notoriedade com o aparecimento de tecnologias como por exemplo, as máquinas fotográficas digitais, *PDA's*, telemóveis, *GPS*, entre outras tecnologias emergentes. O facto é que a evolução acompanha-nos no dia-a-dia, e temos de saber lidar com ela. Daí surge o verdadeiro desafio da preservação digital e das várias estratégias que têm vindo a ser criadas ao longo do tempo com o intuito de acompanhar a evolução, fazendo com que a preservação digital seja conseguida.



Figura 4 – Suportes físicos (Disquete, DC, DVD, Cartão de Memória, USB Flash Drive)

No entanto, o fracasso na preservação destes recursos acarretará à perda irreversível do registo, da prova, do testemunho, da memória [Borbinha et al 2002].

Para além de um formato, o objecto digital também assume estados ao longo do seu percurso, por isso, é também fundamental falar sobre os estados que um objecto digital pode assumir, já que um objecto digital varia de estado consoante o suporte em que se encontra. Um objecto digital pode então assumir quatro estados distintos: o físico, o lógico, o conceptual e o experimentado.

O objecto no estado físico é o objecto que se encontra em suporte no hardware. O objecto em estado lógico é o objecto que se encontra no suporte do software, ou seja, é o objecto depois de ser interpretado pelo respectivo software. O objecto no estado conceptual é o objecto que é disponibilizado ao consumidor, depois de ter sido interpretado pelo software respeitante. Por último temos o objecto experimentado que é a interpretação mental que o consumidor faz do objecto conceptual. Para uma melhor interpretação podemos observar a figura 5 [Ferreira 2006].

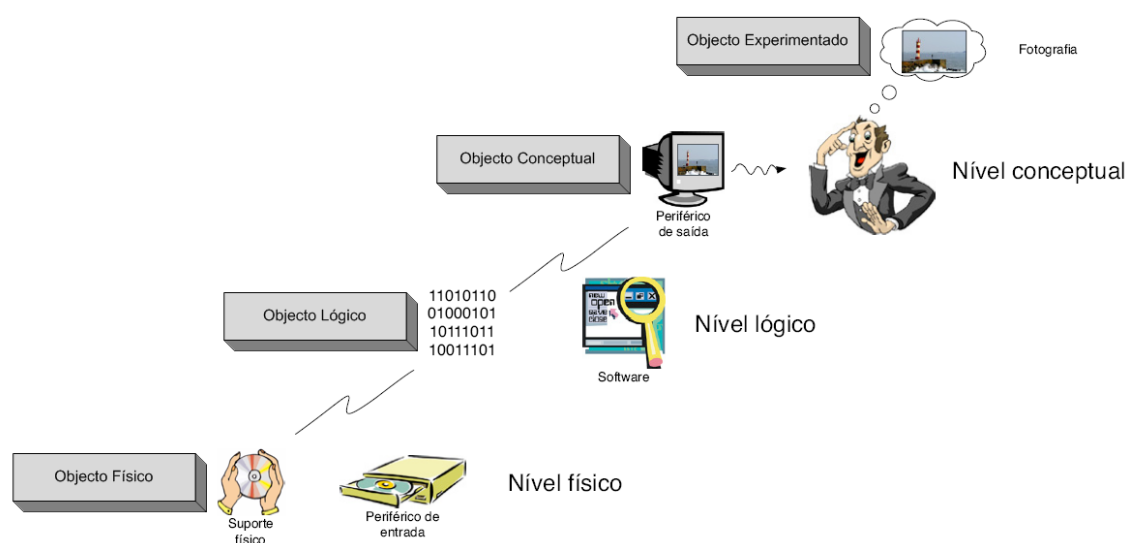


Figura 5 – Estrutura física e lógica de um objecto digital

Estes são os quatro estados que um objecto digital pode assumir ao longo da sua vida, sendo que estão relacionados e que um não existe sem que o seu antecessor tenha existido e assim respectivamente pela ordem que foram mencionados anteriormente. É importante salientar que para um determinado objecto, se um dos níveis se tornar obsoleto o objecto conceptual deixará de estar acessível [Ferreira e Ramalho 2008].

7. Estratégias de Preservação

Para solucionar os problemas inerentes à preservação digital, foram criadas várias estratégias de preservação. As estratégias de preservação digital devem superar as limitações dos recursos digitais, como a fragilidade dos suportes de armazenamento magnéticos ou ópticos [Proença e Lopes 2004]. A informação pode tornar-se inacessível se os suportes que permitem a sua leitura se tornarem obsoletos, havendo uma vulnerabilidade do meio digital. Tanto o hardware como o software se tornam obsoletos, sendo que o software se torna obsoleto mais rapidamente que o hardware.

Ao longo do tempo, foram e continuam a ser criadas várias estratégias que se têm vindo a consolidar, tais como:

- Preservação tecnológica
- Encapsulamento
- Emulação
- Impressão em papel
- Migração

- Rosetta Stone
- Universal Virtual Computer (UVC)

Preservação tecnológica é uma estratégia de preservação digital que consiste na conservação e manutenção de todo o software e hardware, mantendo-o em condições que permitam consultar a informação aí guardada. É de salientar que esta estratégia é bastante dispendiosa.

Encapsulamento é uma estratégia que permite a manutenção do formato original do objecto digital. Mas para fazer com que esta estratégia seja bem sucedida, é também necessário que o objecto digital contenha um conjunto de instruções que permitam interpretar os formatos do ficheiro, bem como o seu conteúdo.

Emulação assenta essencialmente na utilização de um software, ao qual damos o nome de emulador, que seja capaz de reproduzir o comportamento de um determinado hardware e/ou software, numa plataforma com a qual não era compatível.

Isto parte do princípio de que, através da utilização de tecnologias actuais, se possa reproduzir o ambiente original em que foi criado um determinado objecto, como o sistema operativo e o respectivo software no qual foi criado. Para uma melhor aplicação desta estratégia, é aconselhável proceder-se a uma descrição detalhada da tecnologia que foi utilizada para recriar o ambiente.

Esta estratégia assume como princípio fundamental, ser uma solução que permite a preservação de um determinado objecto no seu formato original, garantindo a sua disponibilidade para o consumidor ao longo do tempo.

No entanto, nem sempre é possível aplicar esta estratégia, como por exemplo, no que diz respeito a software, que por vezes se torna difícil de descrever aplicações que sejam capazes de serem reproduzidas posteriormente, este problema torna-se ainda mais acentuado, quando as aplicações em causa são multimédia e hipermédia. Um outro problema da emulação, é o facto dos emuladores se tornarem também obsoletos, fazendo com que se deixe de conseguir reproduzir as aplicações, perdendo-se toda a informação por elas gerada por se encontrar inacessível.

Impressão em papel é uma estratégia que permite a preservação dos objectos no tradicional arquivo, tendo esta várias limitações, como por exemplo, o problema de não permitir a preservação de documentos multimédia, com som, movimento, fórmulas, estruturas de informação, etc. Um outro aspecto que não abona a favor desta estratégia, é o facto do papel ter um ciclo de vida limitado, deteriorando-se ao longo do tempo.

Migração é a estratégia que permite a transferência de material digital de uma plataforma para outra, adaptando os recursos digitais ao ambiente de chegada. É um processo responsável pela reorganização dos elementos de informação quando o software ou hardware se tornarem obsoletos.

O maior objectivo do processo de migração, centra-se sobretudo na preservação dos objectos digitais de forma a que estes sejam compatíveis com as tecnologias actuais.

As grandes desvantagens deste processo colocam-se na possível perda de informação ao nível do conteúdo, estruturas, aparência, da incompatibilidade existente entre formatos de origem e de destino e também da possibilidade de algumas das propriedades dos objectos digitais não serem correctamente transferidas.

O planeamento da migração é difícil de se fazer, pois existe pouca experiência com os vários tipos de migrações que podem ser necessárias para manter o acesso a objectos digitais complexos durante longos períodos de tempo [Hedstrom 1997].

Apesar da estratégia de migração nos ajudar a preservar a informação, esta não é uma preservação permanente, pois com a constante evolução tecnológica o formato de destino está em eminente risco de se tornar obsoleto.

Rosetta Stone foi descoberta pelo exército de Napoleão em 1799 no Egipto. Esta pedra proporcionou aos investigadores obterem o mesmo texto em egípcio demótico, grego e em hieróglifos egípcios. Através destes textos, a *Rosetta Stone* serviu de chave para *Jean – François Champollion* (1822) e a *Thomas Young* (1823) decifrarem os hieróglifos, visto que o grego ainda é uma língua actual e através da confrontação com as restantes foi-lhes permitido decifrar os hieróglifos egípcios, bem como o egípcio demótico.

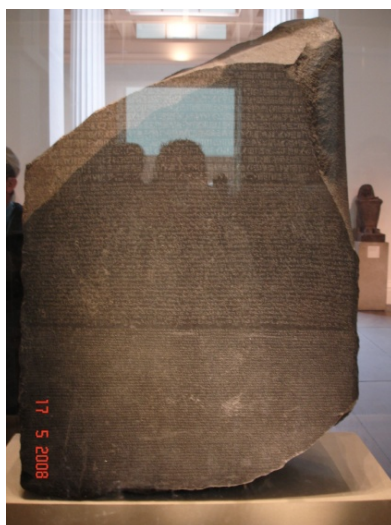


Figura 6 – Rosetta Stone (British Museum 17/05/2008)

Mas afinal em que é que isso se traduz na preservação digital? Em termos de preservação digital, a estratégia baseia-se no mesmo pressuposto, ou seja, dada a amostra de referência, por exemplo, a versão impressa de um documento e as normas de codificação num formato actual, a qualquer momento no futuro podemos criar uma terceira versão da amostra num formato futuro. Ao comparar a amostra actual com a amostra inicial, podemos deduzir as regras para traduzir a partir do formato original para o actual e aplicar estas regras para converter documentos preservados a partir do formato original para um formato actual. É importante para o êxito desta abordagem, construir uma amostra bem caracterizada e extremamente resistente ao longo do tempo [Thibodeau 2005].

Universal Virtual Computer (UVC) é uma estratégia de preservação descrita para facilitar a manutenção dos programas do computador através do tempo.

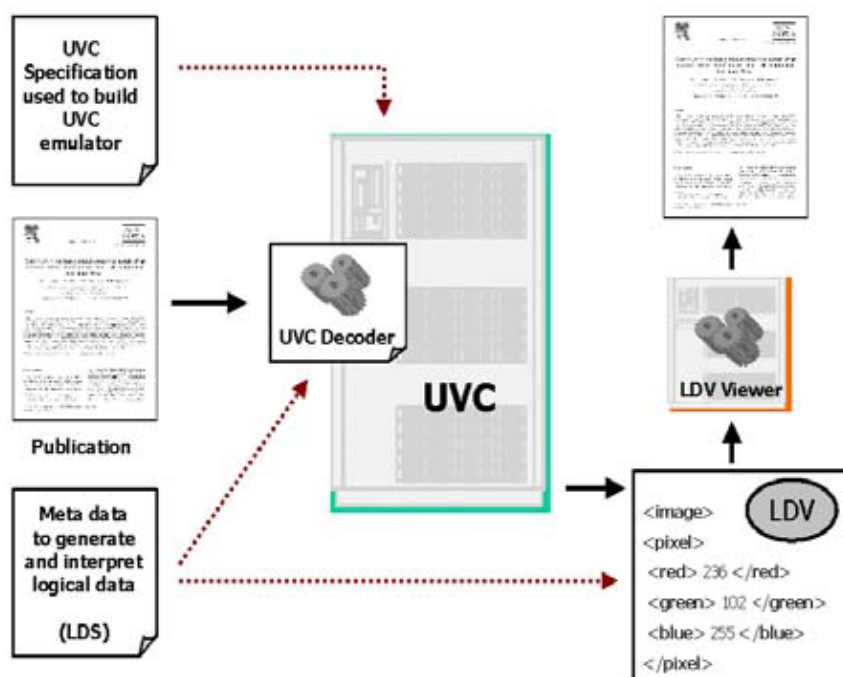


Figura 7 – Modo de Funcionamento do UVC

O *UVC* é baseado em conceitos existentes desde o início da era dos computadores. Por ser virtual, pode ser descrito de uma forma simples e lógica. É esperado que a descrição do *UVC* tenha força suficiente para sustentar programas futuros para construir um emulador *UVC*. Para preservar esta peça crucial de informação, a descrição deverá ser gravada digitalmente, em papel ou em microfilme.

A actual interpretação através do emulador *UVC* deverá permitir que os programas que corram tenham sido escritos em *UVC*. Desta forma, esses programas não são lidos numa instalação, tempo ou ambiente específicos, assegurando a longevidade digital.

Como é que programas podem correr no *UVC*, ajudam a correr publicações digitais que são gravadas nos dias de hoje? Usando um programa específico para traduzir um ficheiro original chamado *Logical Data View (LDV)*.

O *Decoder* tem de ser desenvolvido no presente, sendo guardado juntamente com a publicação original e para o testar, o emulador *UVC* é construído para uma plataforma actual, para que o *LDV* possa ser gerado e testado extensivamente. Assim, a qualidade do *output* é assegurada.

O *Decoder* poderá ser executado no futuro e comportar-se-á exactamente da mesma forma como foi testado, porque também correrá num emulador *UVC*.

O *LDV* poderá ser entendido em qualquer momento do futuro pois é baseado em princípios básicos. De forma a render o *LDV* no futuro, de modo a parecer o objecto digital original, o *viewer* deverá ser construído e capaz de correr em plataformas futuras.

Programadores futuros deverão ser capazes de construir este tipo de *viewer*, por causa da estrutura simples do *LDV* e pelo *Logical Data Schemma (LDS)* que arquivado com o *Decoder* explica a estrutura do *LDV*.

O *LDV* será gerado executando o *Decoder* durante a fase de entrega. De forma a ser capaz de correr este programa, primeiro o emulador *UVC* tem de ser construído para o ambiente actual.

Em seguida, o *LDV* é gerado, e o *viewer* pode ser desenvolvido através das especificações do *LDV* como fora descrito inicialmente no *LDS*.

Porque o *Decoder* trabalhará exactamente como quando fora testado na fase de arquivo, a correcta entrega do *LDV* é garantida [Wijngaarden e Oltmans 2004].

8. A Preservação digital e as Organizações

A preservação digital nas organizações surge como uma necessidade que esta tem em garantir que a sua informação seja mantida e descodificada ao longo dos tempos. Qualquer responsável por uma determinada organização deve ter em mente que a informação gerada pela organização ao longo do tempo é vital para a mesma, visto que, se alguma informação se perder, pode trazer graves problemas para a organização [Gladney 2006].

Podemos dizer que as organizações assumem a preservação digital como uma questão de segurança, ou seja, para uma organização a forma como optam por manter a sua informação em

formato digital assume-se principalmente como uma norma de segurança da própria organização. Mas a verdade é que nem todas as organizações estão correctamente sensibilizadas para a preservação digital, daí muitas organizações procederem à implementação de novos sistemas de informação, sem antes garantirem que a sua informação gerada anteriormente seja interpretada por esses novos sistemas.

Assim, para se proceder a uma implementação de um novo sistema, torna-se necessário ter uma política de preservação digital na organização durante a construção de um repositório de preservação digital de confiança [RLG-OCLC 2002]. É essencial ter uma actividade para minimizar os riscos de perda, de replicação e de estado da informação. Nesta perspectiva assumem-se assim quatro grandes riscos possíveis numa organização [Moore et al 2006]:

- *Technology failure* (Falha tecnológica)
- *Technology evolution* (Evolução tecnológica)
- *Collection evolution* (Evolução dos formatos)
- *Organizational failure* (Falha organizacional)

Abstraindo-nos da amplitude organizacional, e tendo em conta os tipos de riscos, podemos dizer que todos os tipos de riscos são globais, ou seja, são possíveis em qualquer ambiente, excepto os riscos referentes às *Organizational failure*, que são riscos específicos ao ambiente organizacional. Para tal e com o intuito de minimizar estes riscos, foram criados mecanismos, como o *SRB*, que procuram actuar sobre as seguintes categorias [Moore et al 2006]:

- *Checksum Validation*
- *Synchronization*
- *Replication, Backups, and Versions*
- *Federations*

No entanto, as organizações têm que ter presente que depender de *software* e formatos proprietários, que não são compatíveis com as principais normas comuns, pode fazer com que a migração seja mais frequente, daí a necessidade de se adoptar um modelo de preservação da informação [Campos e Saramago 2007].

Todo o tipo de organizações devem ter em conta a preservação da sua informação, sendo as instituições patrimoniais (como por exemplo os museus), aqueles que maior interesse têm na preservação da sua informação e conseqüente preservação digital. No entanto, a preservação digital não é, nem deve ser uma linha de trabalho exógena à missão fundamental da instituição.

O objectivo da preservação digital nesses casos, pode ter mais ou menos a ver com a acessibilidade a longo prazo e menos com a salvaguarda de um objecto enquanto forma física, como acontece com a preservação de património analógico. Porém, não pode ser encarada como uma tarefa marginal, a desempenhar exclusivamente por agentes que não os curadores de património. Cada vez mais, as competências em preservação digital devem ser absorvidas pelas organizações, bem como pelos seus agentes, desenvolvendo-se novas competências ao nível da produção de metadados e da gestão dos repositórios digitais, tendo em conta as metodologias existentes para a implementação de um repositório digital de confiança [Rodrigues 2003].

No enquadramento da preservação digital com as organizações, nomeadamente as institucionais, é necessário proceder-se ao estudo e investigação sobre qual a melhor forma de digitalizar os seus documentos que se encontram em papel e que são um problema sério para que se proceda a uma política de preservação digital.

Existem inúmeras organizações com toneladas de documentos em papel, e neste momento levaria imenso tempo a digitalizar todo esse material de forma a poder guardá-lo num repositório digital. Seria extremamente importante proceder-se a estudos e investigações sobre diferentes formas de digitalização que sejam mais rápidas do que aquelas que temos hoje em dia.

É também de extrema importância que se consiga fazer a digitalização de um documento mais rapidamente, caso contrário as instituições nunca irão proceder a uma política de preservação digital porque além de extremamente dispendiosa, levará demasiado tempo até que se consigam digitalizar todos os documentos. Mas esta seria a maneira de terminar com o caos em termos de documentação, que existe em várias instituições, como por exemplo em tribunais, câmaras municipais, museus, bibliotecas, entre outras.

No entanto, para aquelas organizações que ainda estejam no seu início ou que tenham poucos documentos em papel, estas devem implementar rapidamente uma solução de preservação digital de modo a preservarem a sua informação, em prol do tradicional arquivo de documentos em papel.

A preservação digital torna-se mais fácil e abrangente se as leis estatais forem também elas mais permissivas, de forma a acabar com certas imposições de que determinados documentos têm de estar em papel.

9. Conclusão

Hoje em dia somos confrontados com um aumento exponencial na criação da informação digital, daí a preservação digital ter ganho cada vez mais importância nas nossas vidas. Ao

longo da realização deste trabalho foi possível verificar que a preservação digital é uma questão à qual temos de dar particular atenção, pois esta surge como um meio de garantir o acesso à informação a médio e longo prazo. Para garantir esse mesmo acesso à informação a médio e longo prazo torna-se indispensável a criação de repositórios digitais que sejam capazes de guardar a informação, tendo estes o auxílio de um software que lhes permita um melhor controlo do repositório. No entanto, é necessário ter em mente durante a implementação o modelo de referência OAIS.

Para o sucesso da preservação digital é importante proceder-se à implementação da mesma, tendo como base uma estratégia de preservação que garanta a integridade dos objectos digitais independentemente dos seus formatos e suportes, bem como da evolução dos mesmos. Apesar do crescimento da criação da informação, temos que ter conta que nem toda a informação que é criada tem de ser preservada, cabendo a quem implementa uma solução decidir qual a informação que deve ser preservada.

No decorrer deste trabalho encontrei vários tópicos de investigação futura, sendo que a própria preservação digital no seu todo pode ser encarada como um tópico de investigação. Vendo mais específica e intrinsecamente na preservação digital, julgo que as estratégias de preservação são um alvo de investigação porque ainda há muito que a fazer até se encontrar a estratégia ideal que garanta o sucesso por excelência da preservação digital.

Um outro tópico de investigação diz respeito ao processo de digitalização de documentos, que ainda é bastante lento e em alguns casos impraticável, podendo estar aí uma base de investigação de forma a encontrar-se um método de digitalização mais rápido e eficaz que os métodos actuais.

Os formatos e suportes podem ser também um tópico de investigação na medida em que a evolução tecnológica nos vai forçando à descoberta de novos formatos e suportes de objectos digitais.

Um último tópico de investigação penso que poderá ser centrado nos softwares de gestão dos repositórios digitais, com a finalidade de se obter a máxima performance por parte dos repositórios no acesso à informação.

Para concluir este trabalho posso dizer que este tema foi de bastante agrado, já que este é bastante interessante pois será ainda alvo de muita pesquisa e descoberta no futuro, dada a sua importância para a humanidade. Apesar de toda a dificuldade na sua realização o resultado final acabou por exceder as minhas expectativas.

10. Referências

- Ferreira, M., “Introdução à Preservação Digital: Conceitos, estratégias e actuais consensos”, Escola de Engenharia da Universidade do Minho, 2006, URL: <https://repositorium.sdum.uminho.pt/bitstream/1822/5820/1/livro.pdf> .
- Hitchcock, S., Brody, T., Hey, J. e Carr, L., “Digital Preservation Service Provider Models for Institutional Repositories”, D-Lib Magazine, 2007, URL: <http://www.dlib.org/dlib/may07/hitchcock/05hitchcock.html> .
- Saramago, M., “Metadados para preservação digital e aplicação do modelo OAIS”, Biblioteca do Departamento de Biologia Animal da Faculdade de Ciências da Universidade de Lisboa, 2004, URL: <http://badinfo.apbad.pt/congresso8/comm2.pdf> .
- Moore, R., Rajasekar, A., Wan, M., Schroeder, W., Marciano, R., “On Building Trusted Digital Repositories”, 5th e-Science All Hands Meeting, 2006, URL: <http://www.allhands.org.uk/2006/proceedings/papers/602.pdf> .
- Campos, F., Saramago, M., “Preservação digital de longo prazo em instituições patrimoniais: Reutilização e adaptação de metadados”, 2007, URL: <http://badinfo.apbad.pt/Congresso9/COM42.pdf> .
- Thomaz, K., Soares, A., “A preservação digital e o modelo de referência Open Archival Information System (OAIS)”, Data Grama Zero – Revista de Ciência da Informação, 2004, URL: <http://dici.ibict.br/archive/00000342/> .
- Ferreira, S., “Repositório Institucional em Comunicação: o projeto Reposcom implementado junto à Federação de Bibliotecas Digitais em Ciências da Comunicação”, Enc. Bibli: R. Eletr. Bibliotecon. Ci. Inf., Florianópolis, 2007, URL: http://www.encontros-bibli.ufsc.br/bibesp/esp_06/bibesp_esp_06_pintoferreira_esp_20071.pdf .
- Baptista, A., Costa, S., Kuramoto, H., Rodrigues, E., “Comunicação Científica: o papel da Open Archives Initiative no contexto do Acesso Livre”, Enc. Bibli: R. Eletr. Bibliotecon. Ci. Inf., Florianópolis, 2007, URL: http://www.encontros-bibli.ufsc.br/bibesp/esp_06/bibesp_esp_06_baptista_esp_20071.pdf .
- Proença, A., Lopes, S., “Digital Preservation”, Universidade da Beira Interior, Departamento de Letras, 2004, URL: http://www.di.ubi.pt/~api/digital_preservation.pdf .
- Hedstrom, M., “Digital Preservation: A Time Bomb for Digital Libraries”, Computers and the Humanities, 1997, URL: <http://www.springerlink.com/content/h73v57h6587k417n/> .
- Hedstrom, M., Montgomery, S., “Digital Preservation Needs and Requirements in RLG Member Institutions”, Research Libraries Group, 1998, URL: <http://www.oclc.org/programs/ourwork/past/digpresneeds/digpres.pdf> .
- Wijngaarden, H., Oltmans, E., “Digital Preservation and Permanent Access: The UVC for Images”, Koninklijke Bibliotheek, National Library of the Netherlands, 2004, URL: http://www.kb.nl/hrd/dd/dd_links_en_publicaties/publicaties/uvc-ist.pdf .
- Gladney, H., “Principles for Digital Preservation: Focusing on end users’ needs rather than those of archiving institutions”, Communications of the ACM, 2006, URL: <http://portal.acm.org/citation.cfm?doid=1113034.1113038> .
- Ferreira, M., Ramalho, J., “Preservação Digital: das normas às boas práticas”, Workshop sobre representação digital de objectos com baixo relevo usando PTM’s, 2008, URL: <http://repositorium.sdum.uminho.pt/dspace/handle/1822/7760> .

- Borbinha, J., Henriques, C., Serqueira, J., Lopes, B., “Manifesto para a Preservação Digital”, Cadernos BAD, 2002, Acedido a 03/05/2008, URL: <http://www.bn.pt/agenda/ecpa/manifesto.html> .
- Thomaz, K., “A Preservação de Documentos Eletrônicos de Carater Arquivístico: novos desafios velhos problemas”, Escola de Ciência e Informação da UFMG, 2004, URL: http://www.bibliotecadigital.ufmg.br/dspace/bitstream/1843/VALA-68ZRKF/1/doutorado_katia_de_padua_thomaz.pdf .
- CSSDS, “Reference Model for an Open Archival Information System (OAIS)”, Consultative Committee for Space Data Systems, 2002, URL: <http://public.ccsds.org/publications/archive/650x0b1.pdf> .
- Maniatis, P., Roussopoulos, M., Giuli, T., Rosenthal, D., Baker, M., “The LOCKSS Peer-to-Peer Digital Preservation System”, ACM Transactions on Computer Systems, 2005, URL: <http://berkeley.intel-research.net/maniatis/publications/TOCS2005.pdf> .
- RLG-OCLC, “Trusted Digital Repositories: Attributes and Responsibilities”, Research Libraries Group, 2002, URL: <http://www.oclc.org/programs/ourwork/past/trustedrep/repositories.pdf> .
- Rodrigues, M., “Preservação Digital de Longo Prazo: Estado da arte e boas práticas em repositórios digitais”, Departamento de Ciências e Tecnologias de Informação, ISCTE, 2003, URL: http://dited.bn.pt/8927/37/Tese_Preservacao_Digital_Lurdes_Saramago.pdf .
- Thibodeau, K., “Overview of Technological Approaches to Digital Preservation and Challenges in Coming Years”, Conference Proceedings, 2005, Acedido a 07/05/2008, URL: <http://www.clir.org/pubs/reports/pub107/thibodeau.html> .
- Registry of Open Access Repositories (ROAR), “<http://roar.eprints.org/>”, Acedido a 12/05/2008.
- Registry of Open Access Repository Material Archiving Policies (ROARMAP), “<http://www.eprints.org/openaccess/policysignup/>”, Acedido a 12/05/2008.
- Open Directory of Open Access Repositories (OpenDOAR), “<http://www.opendoar.org/>”, Acedido a 12/05/2008.