

Preservação de conteúdos digitais audiovisuais

Os registos audiovisuais (AV) do século XX estão em risco e a digitalização tem sido uma solução, o que cria um novo problema: a preservação dos registos digitais audiovisuais. Estes ficheiros têm requisitos (tamanhos, formatos específicos) que não estão devidamente adequados à tecnologia disponível. Ainda que as melhores práticas sejam recomendadas, são necessárias três grandes alterações: 1) as colecções dos registos audiovisuais devem utilizar as tecnologias da biblioteca digital e da preservação digital; 2) a tecnologia deverá desenvolver-se de forma a suportar a evolução dos meios de comunicação ao longo dos tempos; 3) todo o armazenamento e a tecnologia de informação, em geral, deverão evoluir de forma a suportarem os requisitos específicos dos registos AV.

Problemas da preservação do audiovisual digital

O acervo audiovisual europeu (em arquivos ou outras colecções formais) está estimado em cerca de 50 milhões de horas de áudio, vídeo e filme, a maioria em formato analógico. Cerca de 70 por cento deste material corre risco agora, e todo ele estará em risco dentro de 30 anos – devido à obsolescência, à deterioração e à desactualização dos formatos (obsoletos).

Iniciaram-se importantes programas de digitalização: estima-se que cerca de 10 milhões de horas já foram digitalizadas na última década. Enquanto se ocuparam das transferências dos conteúdos de cassetes, bandas magnéticas e discos – como uma solução de preservação – os restantes aperceberam-se que os ficheiros digitais apresentam os seus próprios problemas de preservação.

As grandes colecções de ficheiros digitais apresentam um problema de gestão técnica; a solução é a tecnologia de bibliotecas digitais. Os ficheiros precisam de manutenção: têm que ser nomeados, frequentemente transferidos para um novo suporte de armazenamento, copiados para acesso, codificados conforme as alterações das necessidades de acesso e submetidos a testes de validação. Precisam da acção dos seus metadados, variam da catalogação à selecção automática (para um acesso normalizado e global). A manutenção manual é pura e simplesmente impossível – e incorre em demasiados erros – sobretudo quando as colecções atingem determinado tamanho. A tecnologia das bibliotecas digitais fornece ferramentas de automatização para a criação, para a manutenção e para as necessidades de manipulação e de acesso às grandes colecções de ficheiros. Existem muitos guias orientadores sobre as tecnologias das bibliotecas digitais.

Dois mundos: A tecnologia das bibliotecas digitais provém do mundo das bibliotecas académicas. As colecções AV estão largamente fora deste mundo. Os maiores produtores de conteúdos são as emissoras, outras participações importantes estão nos museus de cinema bem como noutras instituições culturais e patrimoniais (uma das maiores colecções de películas do Reino Unido encontra-se Museu Imperial da Guerra).

As emissoras variam mas é comum que as equipas informáticas, bem como aqueles que decidem e financiam as soluções técnicas não conheçam nada de bibliotecas académicas nem de tecnologia de bibliotecas digitais.

O primeiro obstáculo perante a necessidade de preservar ficheiros AV é conhecer, compreender, financiar e utilizar as ferramentas já existentes das bibliotecas digitais que permitam alterar uma pilha de ficheiros numa colecção organizada.

O segundo obstáculo é reconhecer que as ferramentas das bibliotecas digitais permitem uma gestão (os ficheiros tornam-se acessíveis e não se perdem) mas não abrangem a preservação. Os ficheiros enfrentam várias questões relacionadas com a obsolescência, designada por tecnologia de preservação digital – métodos para assegurar os ficheiros em vias de se tornarem obsoletos possam ser migrados conforme novas normas e novos formatos, métodos para emulação de ambientes IT desactualizados que permitam o prolongar da vida dos formatos obsoletos, de critérios para avaliar a fiabilidade de um repositório digital e ainda uma metodologia global: OAIS .

É difícil encontrar nas equipas informáticas que se ocupam das colecções AV quem tenha ouvido falar do OAIS, o que estabelece limites para o seu financiamento e implementação. Felizmente, o Projecto Europeu (EC) MEMORIES está a desenvolver o OAIS e procedimentos relacionados especificamente para colecções áudio e vídeo.

O terceiro obstáculo está relacionado com as necessidades específicas dos ficheiros AV, que não são completamente suportados pelas tecnologias das bibliotecas digitais nem pelas de preservação digital, como se discute a seguir.

Referências e informações adicionais

Inquéritos sobre o Audiovisual
(Audiovisual Status Surveys)

PRESTO : <http://presto.joanneum.ac> at
PrestoSPACE:

<http://www.prestospace.eu>

TAPE: [http://tape-](http://tape-online.net/survey.html)

[online.net/survey.html](http://tape-online.net/survey.html)

[Http://tape-](http://tape-online.net/docs/audiovisual_research_collections.pdf)
[online.net/docs/audiovisual_research_collections.pdf](http://tape-online.net/docs/audiovisual_research_collections.pdf)

Guias Gerais para Preservação

<http://www.bbcarchive.org.uk>

<http://digitalpreservation.ssl.co.uk>

Lista geral de ferramentas de curadoria

<http://twiki.dcc.rl.ac.uk/bin/view/Main/DevelopmentToolList>

JHOVE: <http://hul.harvard.edu/jhove/>
DROID:

<http://droid.sourceforge.net/wiki/index.php/Introduction>

PRONOM:

<http://www.nationalarchives.gov.uk/pronom>

National Library of New Zeland Library
metadata extractor:

<http://meta-extractor.sourceforge.net/>

OAI:

<http://www.openarchives.org/OAI/openarchivesprotocol.htm>

OAIS:

<http://dpeonline.org/docs/lavoie:OAIS.pdf>

Migração:

<http://www.library.cornell.edu/iris/migration/>

Emulação:

<http://www.dlib.org/dlib/october00/granger/10granger.html>

Repositório de referência de critério de
avaliação

<http://journals.tdl.org/jodi/article/view/199/180>

Projectos e Iniciativas

MEMORIES: <http://memoriesproject.eu/>

European Digital Library :

<http://europeana.eu>

Formatos MXF :

<http://www.digitalpreservation.gov/formats/fdd/fdd000013.shtml>

Especificação formato WAV

<http://www.digitalpreservation.gov/formats/fdd/fdd000001.shtml>

Apresentações

Video Formatting and Preservation, Carl

Fleischhauer, NDIIPP DLF Forum,

Philadelphia, 6 November 2007:

<http://www.diglib.org/forums/fall2007/presentations/Fleischhauer.pdf>

Problema impar dos dados e ficheiros do audiovisual digital

Devido ao problema dos “dois mundos”, os formatos profissionais de difusão (MXF em particular), não são suportados pela maioria das ferramentas das bibliotecas digitais nem de preservação. Outros formatos “standarizados” são melhor suportados embora muitos (por exemplo AVI, WMV) são protegidos , o que os torna em si mesmos um problema de preservação.

Demais problemas relacionados com o actual conteúdo destes ficheiros.

A maioria dos ficheiros AV é comprimida. Qualquer “qualidade original” existente perdeu-se para sempre durante a compressão. A preservação deverá maximizar a conservação da qualidade; uma possibilidade que precisa de ser definida e incorporada na tecnologia actual.

Conteúdos time-based exigem ferramentas com dimensão temporal (catalogação, pesquisa, edição).

Os ficheiros são complexos. De certa forma, o conceito de pacote foi desenvolvido para reconhecer a complexidade de um típico ficheiro AV: sinais múltiplos, múltiplos tipos de metadados – incluindo subtítulos (time-domain) e numéricos (time – code).

Preservação audiovisual implica a relação de muitos ficheiros interdependentes: codificações com e sem perdas, vários softwares (que suportem os acessos em múltiplos formatos, por exemplo, Real, Windows Media, MPEG, AVI, Quicktime, Flash), várias etapas de edição e de combinações além de uma série de informação sobre direitos: múltiplas partes interessadas, as diferentes entidades, as diferenças de país para país. Deve ser preservado um grande número de informação complexa, que inclua os sinais, os metadados e os direitos.

O acesso As bibliotecas têm uma larga tradição de acesso único: catálogos comuns baseados numa normalização dos metadados, fornece o serviço “qualquer livro, em qualquer lugar”. A maioria das colecções audiovisuais tem a tradição de estar fechada ou apenas aberta ao acesso profissional ou comercial.

As bibliotecas digitais continuam a tradição de expandir e unificar acessos, quer à escala nacional ou internacional, como projecto da Biblioteca Europeia Digital . As colecções AV precisam da tecnologia das bibliotecas digitais para se tornarem acessíveis através de projectos mais amplos como o BED. Por sua vez, a indústria das bibliotecas digitais precisa de se esforçar mais para compreender os problemas dos dados e ficheiros digitais audiovisuais, já aqui apresentados. Em particular, as bibliotecas digitais precisam de ferramentas para acesso data-based tanto para o sinal AV como para os metadados (os direitos, por exemplo, podem variar de um momento para o outro, mesmo num ficheiro AV).

Muitos dos conteúdos AV pertencem a organismos sem tradição de trabalhar com bibliotecas e que preferem limitar o acesso aos “seus” conteúdos. Questões de mercado, de promoção e de direitos entram um portal europeu audiovisual . A BED poderá nunca incluir conteúdos da BBC.

O que fazer

Apesar dos problemas podem fazer-se algumas declarações elucidativas sobre a preservação audiovisual:

Preservar o artefacto: manter o “original”, ainda que comprimido. ”Preservar a cadeia de bits ”. Os conteúdos audiovisuais têm uma vantagem: são volumosos mas num relativamente pequeno número de formatos. Podem existir métodos para tratar os bits.

Descodificar sem compressão, então salvaguardar sem comprimir (e ao mesmo tempo guardar o original). É uma exigência para vídeo (100 GB/hora para 625 linhas TV), mas armazenagem é mais cara.

Fortalecer os metadados: a extensão de ficheiros (por exemplo: wav, avi não é suficiente). Existem mais de 50 variantes de codificação registadas com a definição .wav; MPEG -1 e MPEG-2 usando a extensão, mpg. Na melhor das hipóteses, existirão ferramentas de extracção de metadados; de outra forma, recorre-se à documentação e aos testes manuais.

Não está sozinho: utilize o arquivo dos registos – tipo, repositórios de software, plataformas de emulação e guias de preservação enunciados nas referências.

Autor: **Richard Wright**, *Research & Innovation, BBC*

richard.wright@bbc.co.uk

Tradução: **Sofia Pina**, Sofia.opina@gmail.com