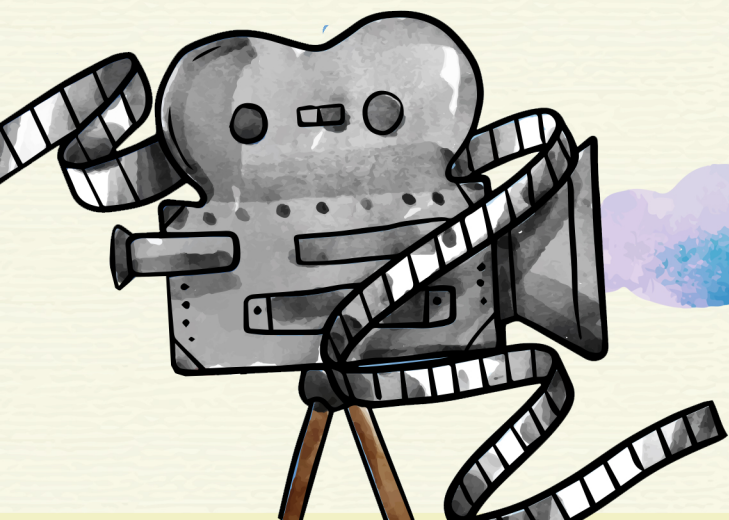


CINEMA NA NUVEM:

A PRESERVAÇÃO E O ACESSO NA ERA DIGITAL



UNIVERSIDADE FEDERAL DE MINAS GERAIS

STEFÂNIA PAULA FERNANDES PEREIRA

**CINEMA NA NUVEM:
A PRESERVAÇÃO E O ACESSO NA ERA DIGITAL**

BELO HORIZONTE
2018

STEFÂNIA PAULA FERNANDES PEREIRA

**CINEMA NA NUVEM:
A PRESERVAÇÃO E O ACESSO NA ERA DIGITAL**

Dissertação de mestrado apresentada ao Programa de Pós-graduação em Artes da Escola de Belas Artes da Universidade Federal de Minas Gerais, como requisito para obtenção de título de Mestre em Artes.

Linha de pesquisa: Preservação do Patrimônio Cultural

Orientadora: Professora Dra. Magali Melleu Sehn

BELO HORIZONTE

2018

Ficha catalográfica

(Biblioteca da Escola de Belas Artes da UFMG)

Pereira, Stefânia Paula Fernandes, 1991-
Cinema na nuvem [manuscrito] : preservação e o acesso na era digital /
Stefânia Paula Fernandes Pereira. – 2018.
132 f. : il.

Orientador: Magali Melleu Sehn.

Dissertação (mestrado) – Universidade Federal de Minas Gerais, Escola de Belas Artes.

1. Cinema – Conservação – Teses. 2. Processamento de imagens – Técnicas digitais – Teses. 3. Preservação pela digitalização – Teses I. Sehn, Magali Melleu, 1961- II. Universidade Federal de Minas Gerais. Escola de Belas Artes. IV. Título.

CDD 702.88

Assinatura da Banca Examinadora na Defesa de Dissertação da aluna **STEFÂNIA PAULA FERNANDES PEREIRA** Número de Registro **2016658155**.

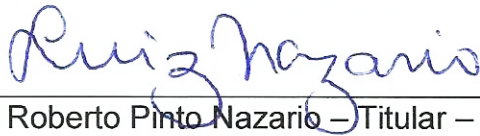
Titulo: "Cinema na nuvem: a preservação e o acesso na era digital"



Profa. Dra. Magali Melleu Sehn – Orientador – EBA/UFMG



Prof. Dr. Carlos Henrique Rezende Falci – Titular – EBA/UFMG



Prof. Dr. Luiz Roberto Pinto Nazario – Titular – EBA/UFMG

Belo Horizonte, 11 de abril de 2018.

AGRADECIMENTOS

Magali Melleu – UFMG, pela orientação e caminho.

Maurício Gino, Luiz Nazário e Leonardo Vidigal – UFMG, pela disponibilidade e auxílio em minha banca qualificação sem mesmo me conhecerem.

Alice Gontijo – UFMG, pelo carinho e ajuda nas dúvidas do percurso, das mais complexas às mais triviais.

Carlos Falci – UFMG, pelas tantas conversas no corredor e indicações sempre incríveis de leituras, além da presença na banca final.

Hernani Heffner – Cinemateca MAM/RJ, pelo acolhimento e ajuda sem reservas e por tempo indeterminado na cinemateca.

Inês Aisengart – Cinemateca Brasileira, pelos muitos gigas de informação compartilhados em prol da pesquisa apresentada nesta dissertação.

Marcella Furtado – CRAV/MIS BH, por, desde a graduação, me inspirar o desejo de conhecimento da preservação audiovisual e deixar as portas do MIS sempre abertas à pesquisa.

Matheus Nagime – ABPA, por formalizar e acolher minha participação na Associação Brasileira de Preservação Audiovisual. Seguimos juntos na busca por parâmetros para a preservação digital.

Ana Lúcia e Francisco – pais que não descuidam, estão sempre achegados e com amor se interessam por minhas decisões.

Clara e Sussu – as melhores irmãs que alguém poderia ter.

Filipe – pelo carinho, paciência e amor.

Marcos e Amanda – Detalhes, por não somente entenderem, mas incentivarem o mestrado, a docência e minha vida acadêmica.

Nathália Cavalieri – Hotmart, por todas as folgas concedidas em benefício deste trabalho, elas foram fundamentais.

Aos amigos e colegas da Detalhes, UNA e Hotmart – pelas tantas conversas, além do interesse e carinho em me ouvir falar da pesquisa.

“As obras audiovisuais não são produzidas no vácuo, mas em uma época e lugar concretos. Como tais, só poderão ser apreciadas corretamente em seu devido contexto”.

Ray Edmondson

RESUMO

Esta dissertação propõe-se a investigar, revisar e contribuir com a metodologia atual de preservação de arquivos audiovisuais digitais, pois sabe-se que estes são frágeis por sua natureza, sejam eles nativos ou convertidos. As cinematecas, arquivos e museus audiovisuais brasileiros carecem de uma construção de critérios que sejam confiáveis, eficazes e, acima de tudo, sustentáveis - dada a realidade econômica do país, bem como a equipe profissional disponível. Para tanto, conhecer o histórico do que chamamos de preservação audiovisual se faz relevante num contexto de pleno crescimento da demanda por armazenamento de arquivos, mesmo os pessoais, momento este capaz de despertar uma consciência sobre a importância geral de "guardar" e "proteger" o que é importante. Faz-se necessário, acima de tudo, compreender a mentalidade e conhecer os hábitos dos usuários desses arquivos: os espectadores dos filmes e vídeos digitais, principalmente online. A web se transformou neste espaço viável, com grandes detentores de conteúdo audiovisual, mas com um armazenamento ainda impensado. Esta dissertação lança um olhar mais aberto à própria internet e a trata menos como vilã e mais como parceira nos processos de preservação audiovisual. Propõe, também, que pessoas e acervos cogitem soluções de armazenamento sem deixar de pensar no acesso e na experiência do usuário. Por fim, sugere alternativas que seriam úteis principalmente no caso dos arquivos audiovisuais públicos brasileiros, sem tantos recursos e com uma demanda crescente.

Palavras-chave: audiovisual, digital, preservação, conservação, acesso, exibição.

ABSTRACT

The object of this dissertation is to investigate, review, and contribute to the current methodology of preservation of digital audiovisual files, as it is widely known they are fragile by nature, whether native or converted. Brazilian cinematheques, archives, and audiovisual museums lack a structure of criteria that is reliable, effective, and, above all, sustainable - due to the country's economic reality, as well as the available professional personnel. Therefore, understanding the history of what we call audiovisual preservation is relevant in a context of ample growth in the demand of file storage, even personal ones, which is translated into an awakening of consciousness about the overall importance of "saving" and "protecting" what is important. It is necessary, above all, to understand the mindset and learn about the habits the users of these files have: the spectators of digital films and videos, especially online. The world wide web has transformed into this viable space, with major holders of audiovisual content, but with its storage still unthought of. This dissertation offers a more open look on the internet itself, perceiving it less of a villain and more as a partner in the processes of audiovisual preservation. It proposes, also, that people and libraries consider storage solutions but acknowledging user access and experience. Finally, it proposes alternatives that would be useful especially for public Brazilian audiovisual archives, which lack in resources and present an increasing demand.

Keywords: audiovisual, digital, preservation, conservation, access, exhibition.

LISTA DE ABREVIATURAS

ABPA – Associação Brasileira de Preservação Audiovisual

APARSEN – *Alliance for Permanent Access*

CCSDS – *Consultative Comitee for Space Data Systems.*

COPPUL – *Council of Prairie and Pacific University Libraries*

COPTR – *Community Owned digital Preservation Tool Registry*

CRA – Centro de Referência Audiovisual

CRAV – Centro de Referência Audiovisual de Belo Horizonte

DCC – *Digital Curation Centre*

DCH-RP – *Digital Cultural Heritage Roadmap for Preservation*

DPC – *Digital Preservation Coalition*

FIAT / IFTA – *Fédération Internationale des Archives de Télévision / International Federation of Television Archives*

FIAF – *International Federation of Film Archives*

FLOSS – *Free/libre open-source software*

HD – *Hard disc*, em português, disco rígido. Desambiguação para *high definition*, alta definição / qualidade – termo não utilizado nessa pesquisa com esse significado.

INA – *Institut National de l'Audiovisuel*

ISO – *International Organization for Standardization.*

LOCKSS – *Lots Of Copies Keep Stuff Safe*

LTO – *Linear Tape-Open*

MAM – Museu de Arte Moderna

MinC – Ministério da Cultura

MIS – Museu da Imagem e do Som

MoMa – *Museum of Modern Art*

OAIS – *Open Archival Information System.*

SAV – Secretaria do Audiovisual

UNESCO – Organização das Nações Unidas para a Educação, a Ciência e a Cultura

LISTA DE FIGURAS

- FIGURA 1 – Frame estático do filme "*L'arrivée d'un train à la Ciotat*".
- FIGURA 2 – Frame do filme *Cinema Paradiso*.
- FIGURA 3 – Fachada da Cinemateca Francesa, em julho de 2016.
- FIGURA 4 – Reportagem sobre o incêndio em 4 de fevereiro de 2016
- FIGURA 5 – Fachada do MAM e prédio da Cinemateca em setembro de 2016.
- FIGURA 6 – Reserva técnica da Cinemateca do MAM em julho de 2013.
- FIGURA 7 – Reportagem sobre transferência do MIS
- FIGURA 8 – Reportagem sobre MIS
- FIGURA 9 – Infográfico - dados sobre o acesso de internet no Brasil.
- FIGURA 10 – Busca de termos no *Google Trends*.
- FIGURA 11 – Falha na busca de termos no *Google Trends*.
- FIGURA 12 – Falha na busca de termos no *Google Trends*.
- FIGURA 13 – Título de reportagem sobre modelo da Televisão Brasileira.
- FIGURA 14 – Gráfico do questionário de pesquisa: faixa etária.
- FIGURA 15 – Gráfico do questionário de pesquisa: escolaridade.
- FIGURA 16 – Gráfico do questionário de pesquisa: emprego.
- FIGURA 17 – Gráfico do questionário de pesquisa: consumo audiovisual.
- FIGURA 18 – Gráfico do questionário de pesquisa: hábitos de consumo.
- FIGURA 19 – Gráfico do questionário de pesquisa: download de filmes.
- FIGURA 20 – Gráfico do questionário de pesquisa: assinatura de TV paga.
- FIGURA 21 – Gráfico do questionário de pesquisa: assinatura Netflix.
- FIGURA 22 – Gráfico do questionário de pesquisa: filmes no YouTube.
- FIGURA 23 – Gráfico do questionário de pesquisa: filmes por demanda.
- FIGURA 24 – Gráfico do questionário de pesquisa: títulos não encontrados.
- FIGURA 25 – Gráfico do questionário de pesquisa: compra de DVDs e Blurays.
- FIGURA 26 – Gráfico do questionário de pesquisa: importância preservação.
- FIGURA 27 – Gráfico do questionário de pesquisa: produção conteúdo audiovisual.
- FIGURA 28 – Gráfico do questionário de pesquisa: hábitos de postagem.
- FIGURA 29 – Gráfico do questionário de pesquisa: porcentagem redes sociais.
- FIGURA 30 – Gráfico do questionário de pesquisa: backup em nuvem.
- FIGURA 31 – Gráfico do questionário de pesquisa: backup em dispositivos físicos.

- FIGURA 32 – Gráfico do questionário de pesquisa: confiabilidade na web.
- FIGURA 33 – Gráfico do questionário de pesquisa: comparativo de imagens.
- FIGURA 34 – Gráfico do questionário de pesquisa: percepção de qualidade.
- FIGURA 35 – Fluxo de trabalho do INA France.
- FIGURA 36 – Página inicial do Archivemática.
- FIGURA 37 – Tradução da versão simplificada do modelo OAIS.
- FIGURA 38 – Captura de uma das telas online do Sistema Legatum.

SUMÁRIO

INTRODUÇÃO.....	10
Capítulo 1 – A PRESERVAÇÃO DE FILMES.....	16
1.1 Contexto mundial.....	17
1.2 Contexto brasileiro.....	29
1.3 Relevância da memória.....	39
Capítulo 2 – PRECISAMOS FALAR SOBRE O DIGITAL.....	42
2.1 A estrutura do arquivo digital.....	48
2.2 Questões relacionadas à preservação.....	55
2.2.1 O modelo de referência OAIS.....	56
2.2.2 HD, LTO, metadados e possibilidades.....	57
Capítulo 3 – YOUTUBE, NETFLIX E A EXPERIÊNCIA DO USUÁRIO	61
3.1 Youtube.....	64
3.2 Netflix.....	67
3.3 Experiência do usuário.....	70
3.3.1 Metodologia e resultados.....	72
Capítulo 4 – E O QUE PODE SER FEITO?	88
4.1 Dossiê de ações.....	94
CONSIDERAÇÕES FINAIS.....	112
REFERÊNCIAS.....	117
APÊNDICE A – Perguntas para composição do questionário.....	124
ANEXO A – Cartilhas DPC para preservação Digital.....	133

INTRODUÇÃO

Os termos preservação e acesso são correlacionados, principalmente quando falamos de arquivos audiovisuais, afinal, filmes são feitos para serem vistos. Esta dissertação de mestrado intitulada “Cinema na nuvem: a preservação e o acesso na era digital” está direcionada à realidade recente da era digital, mais especificamente à preservação de obras audiovisuais e mantém como objeto de estudo a problemática da preservação de filmes em formato digital — sejam eles nativos ou convertidos¹.

Para tanto, faz-se necessário primeiramente considerar a realidade das instituições que cuidam desse tipo de arquivo. A transição para o digital é fato recente e brusco, não havendo, portanto, um guia de como proceder, ou mesmo parâmetros consolidados. Para iniciar a investigação, duas instituições foram visitadas: o MIS BH – Museu da Imagem e do Som de Belo Horizonte (antigo CRAV) e Cinemateca do Museu de Arte Moderna do Rio de Janeiro; para vivência de ambiente cujas questões servirão de norte para a proposição de alternativas adequadas às realidades encontradas, a serem apresentadas no último capítulo.

O levantamento da situação dos acervos levou em conta, inicialmente, a facilidade de acesso em Belo Horizonte. Porém, o MIS possui pouquíssimos exemplares de filmes em formato digital, o que demandou a procura pela segunda instituição. O critério desta foi a abertura das portas à pesquisa e às possíveis sugestões.

De maneira geral, não se pode afirmar que essas instituições são referência no tratamento dos arquivos audiovisuais digitais, pois, originalmente, foram criadas para o acolhimento de películas, VHSs e mídias analógicas. Em BH, até o momento dessa pesquisa, os poucos filmes digitais — resultantes de um edital da prefeitura — encontram-se em HDs externos, fora de suas condições-padrão de preservação e sem backups em qualquer outro dispositivo ou nuvem. No Rio, os HDs se apresentam em maior quantidade, em melhores condições de armazenamento

¹ *Nativos* são arquivos gravados originalmente no formato digital, binário. *Convertidos* são os arquivos analógicos que sofreram migração para o digital por meio de tecnologias conversoras.

(apesar do clima e local desfavoráveis) e a grande maioria possui backup em outro HD externo. Nenhuma das instituições possui servidor ou tecnologias de armazenamento de dados avançadas até a publicação desta.

Ambos os acervos possuem seus itens catalogados e disponibilizados para fins de pesquisa sem grandes impedimentos, a conversa é outra quando o assunto é ceder essas imagens e sons para qualquer tipo de filme ou produção audiovisual — política compreensível, partindo do pressuposto de que a instituição de guarda não detém os direitos autorais da maioria das obras — apenas armazenam. Isso significa que, para ceder o material, é necessário contato com o autor/detentor dos direitos, pagamento desses direitos e/ou autorização prévia.

Embora outras instituições não estejam sendo abordadas diretamente, tal procedimento para cinematecas e arquivos audiovisuais pode ser padrão. Todos os arquivos possuem (ou pelo menos, deveriam possuir) um estatuto e uma política que rege as questões que podem surgir quando o assunto é acesso do público às obras. Ray Edmondson, em seu livro *Filosofia e princípios da arquivística audiovisual* já enuncia a importância de os arquivos funcionarem de maneira

aberta e responsável, consistente, [com] ética e padrões profissionais, [atentando à] credibilidade e integridade do campo dos arquivos audiovisuais, [criando e observando] políticas escritas, diretrizes e normas como quadro de referência para o seu trabalho. Não se basearão em entendimentos ou transações informais, procedimentos intuitivos, ou organizações de carácter pessoal. Esforçar-se-ão para documentar toda a informação pertinente de forma recuperável de modo a não depender da memória ou conhecimento individual.²

Partindo desse pressuposto estabelecido por ele, a pesquisa apresentada nesta dissertação pode contribuir para a maneira pela qual os arquivos trabalham seus materiais e para a criação de parâmetros para o digital.

Uma vez conhecidas as necessidades das instituições, estudou-se a metodologia utilizada atualmente na salvaguarda desses filmes. Após as leituras, pesquisas e

² EDMONDSON, 2013, p. 45.

diálogo com as instituições, percebeu-se a necessidade de levar em conta a expansão da internet e de plataformas de *streaming*³ e armazenamento via *upload*.⁴

Sendo assim, tal pesquisa se faz relevante principalmente, mas não somente, por ser significativa produção de conteúdo em uma área ainda pouco explorada, sobretudo em âmbito nacional. Por considerar algo que é tão recente e delicado. Atualmente, temos muita informação sendo produzida o tempo todo e com o fácil acesso à internet ela também é transmitida e armazenada de inúmeras maneiras. Os suportes e os formatos que o digital abriga não podem sequer ser todos mencionados neste trabalho, tamanha é a velocidade com que eles mudam e também novos suportes e formatos são inventados.

Importante ressaltar que a preservação de filmes analógicos ainda é bastante rígida com relação aos originais e no digital isso não parece fazer tanto sentido. Pesquisadores afirmam perdas a cada cópia de um arquivo digital, mas perda maior seria não ter uma cópia para exibir para o espectador — cujo conteúdo somente faz sentido a partir da presença dele, em sua experiência e percepção do filme.

Sendo a formação da autora em Cinema e Audiovisual, com pós-graduação em documentário e imagens de arquivo, o trabalho nasceu de um forte interesse pela efemeridade do digital após presenciar inúmeras situações de perda de arquivos: *pendrives* e HDs externos queimados, backups em nuvem que desapareceram, arquivos corrompidos e inacessíveis sem explicação aparente e principalmente: o incêndio da cinemateca brasileira no começo do ano de 2016. Isso porque, aparentemente, a digitalização não resolve o problema principal: nossa incapacidade de evitar que tragédias assim aconteçam — mas, num contexto utópico em que essa digitalização tivesse sido feita, o material seria ainda acessível. Claro, não manteria as características originais, mas grande parte do que importa ao espectador, de fato, estaria lá: imagem, som e contexto — interpretáveis e acessíveis, mantendo, inclusive, as perdas e danos originais, desfazendo-se "apenas" da materialidade da película.

³ Termo em inglês que significa transmitir ou receber dados (no nosso caso, os audiovisuais) através da internet e de maneira contínua.

⁴ Termo em inglês que significa enviar um arquivo pelo computador, subir, colocá-lo em outro lugar.

Não se pode tratar com tanta simplicidade algo tão sério e relevante, e também **não se defende aqui, em hipótese alguma, o descarte do original**. Porém, os dados deste trabalho levam também em conta a **experiência do usuário** — porque filmes foram feitos para serem assistidos. Somente guardados eles não produzem efeitos e não servem de nada.

Apesar de ser um trabalho na linha de “Preservação do Patrimônio Cultural”, expandem-se as áreas de busca de informação também para os autores do próprio cinema; da ciência da informação com armazenamento e gerenciamento de dados; dos que se dedicam a preservar arte digital, videoarte e formatos atípicos; bem como os autores da linguística e arquivologia que se desdobram sobre os conceitos de arquivo, biblioteca, colecionismo, dentre outros. Apropriados da maneira correta, todos esses campos e autores fomentam e aprimoram a discussão sobre o tema.

Por se tratar de um tema extremamente contemporâneo, novos autores e produções pouco conhecidas também serão contempladas. Cabe ressaltar que as sugestões metodológicas que aqui serão propostas fogem um pouco do que temos hoje como diretrizes da preservação audiovisual, mas bem argumentadas no que diz respeito ao acesso e à experiência do usuário – mantendo de pé a função básica da preservação.

A pretensão da autora é sustentar os objetivos propostos, que previam analisar as teorias que regem o processo de preservação de acervos audiovisuais digitais. Pretendia-se catalogar os estudos, ações e medidas realizados até então pelas diversas instituições de guarda, bem como a eficácia de tais procedimentos. Para tanto, a dedicação foi larga ao pesquisar teorias existentes no campo da preservação audiovisual; investigando qual o tratamento dado aos arquivos digitais; catalogando os procedimentos vigentes; observando os acervos e procedimentos das instituições para, por fim, propor ideias que cooperem com a área.

Tendo isso em vista, fica mais fácil compreender porque a autora não se dedicou a apenas estudo de caso de algumas obras fílmicas. Isso esvaziaria a proposta, que

poderia ser entendida como metodologia genérica e pouco aplicável. A intenção desta dissertação, no entanto, não é pensar ou propor metodologia oportuna e cabível a todos os casos, mas a compreensão de que existem possibilidades para os acervos e também os arquivos pessoais, levando em consideração o comportamento do público consumidor de filmes e séries, pensando o tratamento e conversão dos arquivos de maneira que garanta o acesso para esse mesmo público.

Dessa forma, propõe-se uma tentativa de acolhimento (por parte das instituições) de novos recursos. Isso amplia as discussões já existentes, como as que a própria Associação Brasileira de Preservação Audiovisual (ABPA) tem promovido. Recentemente, por meio de uma lista de *e-mails*, a autora foi informada de que a padronização de diretrizes para o audiovisual digital ainda não foi finalizada, sendo, portanto, de extrema importância a publicação desse conteúdo e de estudos acadêmicos nesse contexto, como forma de diálogo e auxílio aos que também se debruçam sobre o digital. Esse relatório é, portanto, o começo de uma proposta de recursos para os arquivistas e/ou instituições que trabalham com esse tipo de arquivo e que, por vezes, amedrontados pelas perdas ou pela própria obsolescência, permanecem engessados em seus trabalhos.

A aproximação do objeto de estudo se deu, inicialmente, pela revisão dos conteúdos disponíveis sobre a preservação audiovisual: analógica e digital. No **primeiro capítulo**, intitulado "**A preservação de filmes**", essa evolução será melhor concebida, vista a partir de uma contextualização histórica e local: partindo das primeiras iniciativas no mundo e vindo ao Brasil, dentro de suas especificidades econômicas e climáticas. Transitar pelo analógico se faz importante nesse primeiro momento, porque ele é precursor de toda a mentalidade sobre a necessidade da preservação audiovisual já consolidada atualmente. Uma vez reconhecida a importância desses arquivos, de se ter pessoal qualificado e condições favoráveis ao seu armazenamento, seremos capazes também de pensarmos o audiovisual em sua versão digital, mais recente e repleta de questões e dilemas.

No **segundo capítulo**, torna-se clara a premissa de que "**precisamos falar sobre o digital**". Sabemos que ainda estamos engatinhando no que diz respeito à compreensão de sua estrutura para proposição de meios de preservação eficazes,

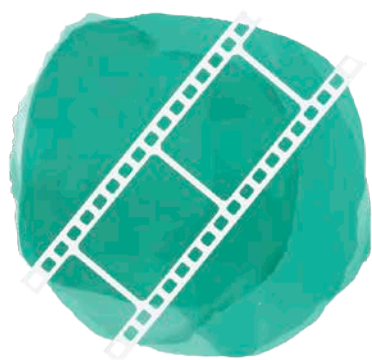
parece complexo falar sobre dados, “zeros e uns”⁵ e tecnologias que amanhã mesmo serão obsoletas. Mais relevante ainda é considerar que suas perdas são quase sempre irreparáveis. O exemplo é clássico: no analógico, o filme é passível de perdas por queima de frames, rasgos, resíduos que arranham ou corroem, dentre outros — mas um frame ou outro ainda são medidas insignificantes, visto o todo do filme que se mantém, mesmo nesses casos. Já no digital, se o arquivo de filme, porventura, se corrompe, perde-se tudo, de uma só vez, sem deixar rastros.

Dessa forma, duas plataformas digitais de armazenamento de conteúdo audiovisual foram eleitas como possíveis soluções e que também inauguram uma nova questão, tratada mais a fundo no **terceiro capítulo**, intitulado "**Youtube, Netflix e experiência do usuário**". Relacionada ao acesso e ao armazenamento desses dados na web, via aplicativos: o pioneiro YouTube e também o Netflix. A escolha foi baseada na premissa básica do acesso, uma vez que ambos se adaptam à velocidade da conexão do usuário e o segundo se adapta também à resolução da tela utilizada. Com isso, pretende-se tratar a própria internet menos como vilã e mais como parceira nesse processo de preservação, principalmente no que diz respeito ao acesso por parte do espectador — para quem se preservam os filmes, afinal.

Nota-se, então, que digital vem trazendo mudanças não somente nos meios de produção e pós-produção de vídeos e filmes, mas também no modo de consumo dos espectadores. Não teremos, portanto, estudo de caso. Uma vez que nem o MIS, nem a Cinemateca possuem (até a data da realização desta dissertação) um plano efetivo de trabalho com seus filmes digitais.

Por fim, no **quarto capítulo**, intitulado "**E o que pode ser feito?**" faremos uma análise baseada nas leituras mais recentes sobre o assunto, seguida de sugestões de metodologias e plataformas, por meio de um dossiê com algumas iniciativas já executadas em diversos lugares.

⁵ Expressão referente aos códigos binários do digital. Segundo definição: "1 bit (binary digit ou dígito binário em português) é a menor unidade de informação que pode ser armazenada ou partilhada e só pode ter 2 valores: 0 ou 1". Fonte: <https://www.significados.com.br/binario/>. Acesso em: 20 nov 2017.



CAPÍTULO 1

A PRESERVAÇÃO DE FILMES

O presente capítulo apresenta a temática da preservação audiovisual — excluindo as técnicas, porque já há uma infinidade de trabalhos nacionais, traduções de métodos internacionais e excelentes manuais teóricos e práticos sobre todo material analógico e películas. As instituições nacionais hoje já trabalham com conhecimento bastante apurado das técnicas e dentro das próprias realidades. Trata-se, antes, de uma contextualização histórica que esclarece quais foram os caminhos pelos quais o próprio cinema teve de passar para se consolidar como forma de arte, e futuramente, como algo que deve ser preservado — mundialmente e nacionalmente. Citando um grande teórico da área e utilizando outros termos:

"a preservação consciente e objetiva da memória é um ato intrinsecamente político e pleno de valores. Não há poder político sem controle sobre os arquivos e sobre a memória. A democratização concreta sempre poderá ser medida à luz deste critério essencial: o grau de participação na formação dos arquivos, o acesso a eles e a participação em sua interpretação"⁶.

Dessa forma, prossigamos. Com a consciência nem sempre clara dessa importância dos arquivos, da própria memória, dos registros de épocas e povos.

1.1 Contexto mundial

Os relatos do que foram os primeiros filmes do mundo são bastante conhecidos, principalmente entre os entusiastas do cinema. Após a criação de vários brinquedos óticos, do domínio das técnicas de fotografia e experiências com imagens congeladas postas em movimento; os irmãos Lumière, pioneiros no cinema, gravaram a chegada de um trem à estação de *La Ciotat* na França⁷ (FIGURA 1). Esse filme representa um grande marco, no qual a história diz que se instaurou o que chamamos de Primeiro Cinema⁸. Depois disso, a produção nunca mais parou, fazendo-se necessário e importante ressaltar que esse mesmo cinema passou por várias transformações ao longo de seus anos de história e que há quem considere⁹ que ainda estamos "engatinhado" nessa forma de arte.

⁶ EDMONDSON, Ray. 2013, p. 42

⁷ *L'arrivée d'un train à la Ciotat* - Auguste and Louis Lumière, França, 1895.

⁸ Período inicial do cinema, até 1930, no qual se tem apenas aproximadamente 30% preservado.

⁹ CICCARINI, Rafael. Centro Universitário UNA - aula. Anotações da autora, 2012.

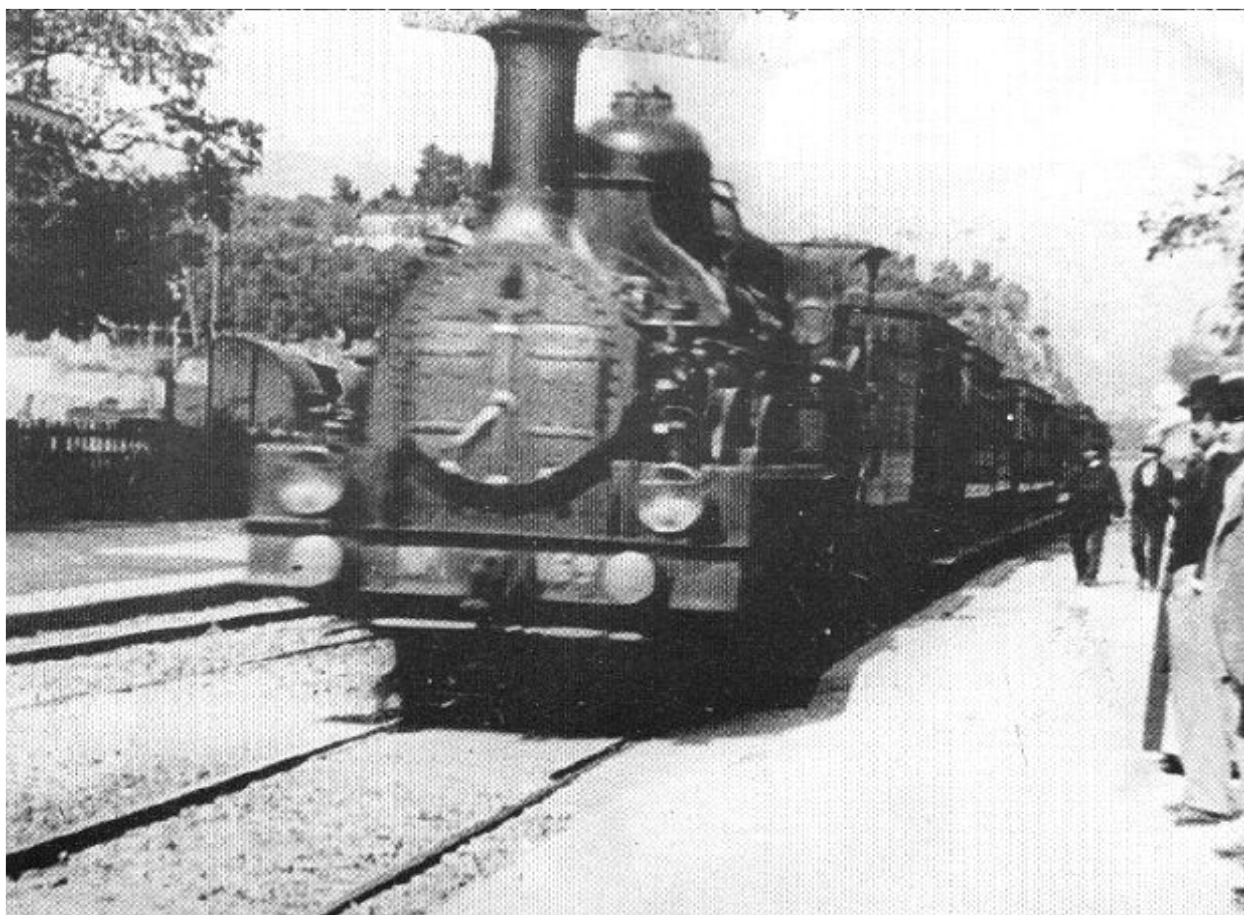


FIGURA 1 – Frame estático do filme *L'arrivée d'un train à la Ciotat*.

FONTE: <https://drnorth.files.wordpress.com/2009/11/larrive-dun-train-a-la-ciotat1.jpg>

Fernando Mascarello (2006), em seu livro-coletânea *História do Cinema Mundial*, aponta-nos um problema recorrente: "há carência, em Língua Portuguesa, de textos introdutórios atualizados e suficientemente amplos"¹⁰ para o estudo do cinema, em outras palavras, falta profundidade. O mesmo ocorre com os textos relacionados à área da preservação audiovisual — o que nos exorta a recorrer constantemente aos autores internacionais.

¹⁰ MASCARELLO, Fernando. 2006, p. 11.

Trataremos o cinema, neste momento, a partir do conceito de Mascarello (2006), cuja inauguração no século XX representou uma "era de predominância das imagens [que] não possuíam um código próprio e estava misturado a outras formas culturais"¹¹. Ora, devemos considerar as ricas misturas de sons e imagens, além das influências do teatro, da dança, da música, pinturas e da própria fotografia. O autor ainda utiliza a expressão "transformação constante"¹² para caracterizar os primeiros vinte anos do cinema, embora saibamos que tais metamorfoses são constantes ainda hoje.

Para ele, "a história do cinema faz parte de uma história mais ampla, que engloba não apenas a história das práticas de projeção das imagens, mas também a dos divertimentos populares"¹³, ou seja, já naquela época era possível pensar o cinema a partir da *experiência do usuário* — conceito que será melhor trabalhado no terceiro capítulo desta dissertação. Ainda em seus primórdios, o cinema se construiu de pequenos avanços tecnológicos: como o surgimento da própria lâmpada, dos materiais tipo filme transparentes (chamado celuloide) e em relação à construção, cada vez mais eficaz, de equipamentos projetores¹⁴. Ainda hoje somos capazes de perceber que a linguagem cinematográfica, bem como os recursos narrativos se esbaldam nos avanços incessantes da tecnologia.

Promover esse retorno aos primórdios cinematográficos se faz necessário porque "durante muito tempo, o cinema dos primeiros 20 anos foi considerado de pouco interesse (...), como apenas um conjunto de desajeitadas tentativas de se chegar a uma forma de narrativa intrínseca ao meio, que se estabeleceria depois"¹⁵. Lentamente, as produções superaram essas limitações iniciais e foram tomando formato específico, dotado de linguagem própria. Apenas na década de 70, os pesquisadores começaram a questionar os critérios usados para a definição do primeiro cinema e dos primeiros filmes, recorrendo aos arquivos e cinematecas para criação de uma nova concepção histórica sobre eles¹⁶.

¹¹ MASCARELLO, Fernando. 2006, p. 17.

¹² MASCARELLO, Fernando. 2006, p. 17.

¹³ MASCARELLO, Fernando. 2006, p. 17-18.

¹⁴ MASCARELLO, Fernando. 2006, p. 18.

¹⁵ MASCARELLO, Fernando. 2006, p. 22.

¹⁶ MASCARELLO, Fernando. 2006, p. 22.

O rigoroso trabalho de pesquisa de estudiosos (...) funcionou, por sua vez, como estímulo à pesquisa sistemática em arquivos e documentos de época, bem como à análise detalhada e cuidadosa das cópias dos primeiros filmes, existentes em cinematecas e arquivos, por uma nova geração de historiadores. Procurava-se tomar contato com materiais primários e cópias de filmes, questionando-se as ideias estabelecidas sobre esse cinema que tinham sido formuladas, muitas vezes, com base na memória pessoal e naturalmente imperfeita dos primeiros historiadores¹⁷.

Tais pesquisas foram impulsionadas por arquivistas visionários, decididos pelo propósito de analisar esses primeiros filmes:

em 1978 a conferência 'Cinema 1900-1906' foi patrocinada pela Federação Internacional dos Arquivos de Filmes (FIAF) [na] Inglaterra. [Nela], pesquisadores e arquivistas debateram juntos novos critérios de datação, identificação e interpretação para os filmes (...). Era a primeira vez que se fazia uma discussão sistemática e coletiva sobre os primeiros filmes de um ponto de vista distinto daquele das histórias clássicas do cinema¹⁸.

Esse processo se deu por um desejo de compreensão das diferenças e peculiaridades dos filmes produzidos nestes períodos, o que levanta a questão "da importância de se entender os filmes em seu contexto específico"¹⁹. Dito isso, vale o retorno à epígrafe deste trabalho: "as obras audiovisuais não são produzidas no vácuo, mas em uma época e lugar concretos. Como tais, só poderão ser apreciadas corretamente em seu devido contexto"²⁰. E é esta a principal premissa dessa produção, pensar o audiovisual em seu contexto atual, de forma especial, o processo migratório que vivemos do analógico ao digital.

Na atualidade, um estudante de cinema ou mesmo um leigo admirador pode assistir [pelo próprio YouTube] à famosa chegada do trem na estação, ou a um espetáculo de Georges Méliès²¹ na lua sem se dar conta de que nada disso existiria se o cinema não tivesse sido preservado. Os que hoje se dedicam a estudar e fazer cinema não teriam essa possibilidade se não fossem esses primórdios acessíveis. Inevitável pensar que, para chegar aonde estamos, alguém, lá atrás, algum dia, pensou em preservar.

¹⁷ MASCARELLO, Fernando. 2006, p. 22

¹⁸ MASCARELLO, Fernando. 2006, p. 23.

¹⁹ MASCARELLO, Fernando. 2006, p. 23.

²⁰ EDMONDSON, Ray. 2017, p. 57.

²¹ Cineasta do primeiro cinema. Pioneiro em efeitos, seu maior filme se chama "*Le voyage dans la lune*" e retrata com maestria e bom humor uma viagem à lua e embates espaciais.

Os registros de como isso se deu, pela primeira vez, ainda são incertos, mas algumas pessoas já tinham o hábito de guardar filmes em suas casas, o que seria um primórdio do colecionador. Ora, a figura do colecionador (FIGURA 2) é fundamental para se entender as razões pelas quais a preservação importa ainda hoje. Eles conseguem ver o que de valioso há no que guardam, e garantem que aquilo ultrapasse seu tempo pré-determinado.



FIGURA 2 – Frame do filme *Cinema Paradiso*, em que o garoto Totó, guarda pedaços de película em seu quarto.

FONTE: TORNATORE, Giuseppe. 1990, DVD, cores.

Walter Benjamin, em seu livro, "O anjo da história", afirma que "há muitos tipos de colecionadores, e em cada um deles atua um feixe de impulsos"²². Interessante pensar como impulso aquilo que move o colecionismo. Pode parecer que se trata de um investimento desnecessário, mas foi esse ímpeto de guardar que fez com que (também) o cinema fosse preservado.

²² BENJAMIN, Walter. 2012, p. 126.

A história do cinema não é feita apenas de grandes invenções, aparelhos engenhosos, demanda por novidades e a constante destruição das películas ao sabor dos novos aperfeiçoamentos e técnicas. Existe outra história, paralela a esta, que acaba tornando-se a principal ponte de ligação entre o mundo novo e o antigo, entre o cinema atual e sua memória. Partindo de práticas isoladas, criando acervos pessoais, cineastas realizadores ou apenas cinéfilos apaixonados pela sétima arte criaram o 'insano' hábito de colecionar filmes, já percebendo que suas películas favoritas desapareciam dos cinemas e, num segundo instante, de suas memórias. Eles entenderam o cinema como o melhor registro de seu tempo, além de ser não apenas um divertimento, mas a mais nova das artes²³.

Marcelo La Carreta, em sua tese²⁴, cria o termo “prossumidores nostálgicos” que significa exatamente isso. Esses colecionadores que, por valor sentimental, econômico ou qualquer outro, tiveram a proatividade de guardar algo que posteriormente desapareceu.

O prossumidor, menos um salvador de fato, antes um entusiasta que vive a apontar indícios de que algo merece ser preservado. (...) Ressalta-se que o prossumidor raramente é um técnico. Não dispõe de verba para restauro. Não participa de uma Cinemateca ou Centro de Restauração. Nota-se, porém, que a criação desses lugares partiu de iniciativas de prossumidores em potencial, de pessoas que passaram da passividade para a prática da salvaguarda de fato, extrapolando o mero colecionismo ao fornecer meios e ferramentas para perpetuar a memória audiovisual²⁵.

É importante evidenciar que "o cinema é uma arte essencialmente efêmera, e já nasceu sem maiores preocupações em ser conservada para as gerações futuras"²⁶. O histórico de perdas, principalmente no Primeiro Cinema, auxilia na compreensão da mentalidade contemporânea de preservação audiovisual e passar por ele se faz necessário neste momento.

Há quem diga que a ideia de preservação surgiu juntamente ao próprio cinema, pelo fato de em 1898 o polonês e cineasta Bolesław Matuszewski já ter projetado a concepção de locais próprios para o armazenamento dos filmes, mas pensando

²³ LA CARRETA, Marcelo. 2005 p. 26.

²⁴ Prossumidores Nostálgicos - UFMG, 2012.

²⁵ LA CARRETA, Marcelo. 2012, p. 14-15.

²⁶ LA CARRETA, Marcelo. 2005 p. 24.

apenas na relevância de seus próprios registros²⁷. Tal projeto não foi suficiente para proclamar um ideal de preservação. Ainda no princípio do cinema, os filmes em películas eram frequentemente descartados depois de exibidos. [Eles] “já haviam atingido o público desejado, já tinham feito dinheiro, ninguém imaginava que poderiam ter outra serventia ou aplicação”²⁸.

Na maioria das vezes, elas [as películas] eram vendidas a quilo, e acabavam virando pente, vassoura ou esmalte. Outras vezes, (...) era queimada pelos realizadores para criar efeitos especiais nas novas produções. Os primeiros projetoristas jogavam na rua os filmes já exibidos. Os que restavam eram cuidadosamente queimados fora das salas, a fim de evitar que os mesmos queimassem essas salas no caso de um incêndio generalizado, fato comum numa época em que a eletricidade, ainda uma novidade, era usada sem controle.²⁹

O autor descreve, nesse trecho, que os filmes eram produtos sem qualquer outro tipo de valor, podendo ser desprezados e descartados após o consumo. Podemos, então, perceber que, ao que parece, não existia uma consciência de preservação nem da importância histórica daquelas imagens e sons. O cinema ainda estava longe de seu reconhecimento como forma de expressão artística. Tal “reconhecimento como bem cultural, como registro histórico, mas sobretudo artístico, é o que funda o discurso de sua preservação”³⁰.

Ray Edmondson, teórico renomado na preservação audiovisual, em seu livro “Filosofia e princípios da arquivística audiovisual” (2013) afirma que nos primórdios do primeiro cinema, não era de forma alguma evidente que os sons e as imagens em movimento possuíssem qualquer tipo de valor duradouro³¹. Embora saibamos que a própria invenção do cinema foi “em grande parte, resultado da curiosidade e do empenho científicos, seu rápido desenvolvimento deveu-se à sua explosão como forma de entretenimento popular”³².

²⁷ STECZ, Solange Straube. 2016, p.31.

²⁸ LA CARRETA, Marcelo. 2005, p. 24.

²⁹ LA CARRETA, Marcelo. 2005, p. 24.

³⁰ QUENTAL, José. 2016, p. 36.

³¹ EDMONDSON, Ray. 2013, p. 92.

³² EDMONDSON, Ray. 2013, p. 92.

Consciente de todos esses antecedentes, há mais ainda o que levarmos em conta no pensamento da história da preservação dos arquivos audiovisuais. É fundamental considerar os pormenores individuais e investimentos dos distintos países nessa “história geral” da preservação audiovisual que ainda se encontra fragmentada:

São vários motivos para esse desenvolvimento desigual da área. Entre eles estão as circunstâncias políticas, históricas e econômicas de determinados países e suas indústrias audiovisuais, as realidades climáticas (os materiais audiovisuais deterioram com mais rapidez em zonas tropicais) e as condições culturais.³³

A história da preservação das imagens em movimento abrange, logo, esses pequenos movimentos em prol de seu reconhecimento como bem cultural:

Lentamente, foi ganhando força o discurso que defendia os produtos desta nascente indústria cultural como tendo valor outro para além do — ou mesmo contra — seu valor comercial e uso. Por volta da década de 1930, na Europa e EUA, pioneiros da preservação audiovisual (...) ‘imaginaram’ que o cinema era uma arte e que tinha um passado, (...) e que por isso deveria ser preservado.³⁴

Impossível, então, precisar uma data formal para o surgimento da preservação audiovisual. “Ela surgiu de fontes difusas, (...) em uma ampla variedade de instituições de coleta e conservação (...) como extensão natural do trabalho que realizavam”³⁵. Com o passar do tempo, as atividades tornaram-se, também, reconhecidas pelas organizações internacionais de arquivos gerais, museus e bibliotecas³⁶.

Pouco a pouco, afirmando sua capacidade, atualmente, interpreta-se o cinema como “expressão da cultura dos povos, as imagens em movimento [hoje] são consideradas pelo seu valor educativo, cultural, artístico, científico, histórico, formando parte do patrimônio cultural de todas as nações”³⁷.

³³ EDMONDSON, Ray. 2013, p. 96.

³⁴ QUENTAL, José. 2016, p. 36.

³⁵ EDMONDSON, Ray. 2013, p.91.

³⁶ EDMONDSON, Ray. 2013, p.91.

³⁷ STECZ, Solange Straube. 2016, p.31.

As cinematecas e arquivos só começaram a surgir de forma sistemática no final dos anos 20 do século XX - A Cinemateca Sueca que depois foi seguida pela criação das instituições alemãs, inglesas, norte-americanas e, certamente, pela que mais se projetou, ao longo do século, a francesa (pela ação de figuras como Georges Franju, Henri Langlois, Jean Mitry e outros) e outras³⁸.

Alguns museus e bibliotecas também adequaram seus acervos para receberem material fílmico. O MoMa, por exemplo, consiste em uma instituição empenhada em salvaguardar a arte moderna e, com o tempo, tornou-se um museu da história de todas as artes, inclusive, do cinema³⁹.

Ele [o MoMa] foi o “responsável pela preservação de pérolas do cinema mudo como *Nanuque, o Esquimó* (*Nanook of the North*, 1922), de Robert Flaherty, e *Intolerância* (*Intolerance*, 1916), de David Wark Griffith, depositados na instituição pelos seus próprios realizadores”⁴⁰.

O nome de Henri Langlois certamente também será encontrado em qualquer pesquisa sobre o assunto, dado que ele foi “um dos mais importantes defensores da memória do cinema”⁴¹. Seu gosto por colecionar os filmes que mais gostava virou costume, o costume virou compulsão e logo depois ele percebeu “que certos títulos de sua coleção tornavam-se ao longo dos anos a única cópia que deles restava”⁴².

Langlois guardava suas raridades em nitrato na banheira da casa, para o desgosto da família, que não podia tomar banho na banheira e ainda corria (sem saber) sérios riscos, pois o aquecimento do banheiro era a gás⁴³. Gastou todo o dinheiro que possuía para adquirir quantos filmes fossem possíveis, até que o governo francês, muito tempo depois, reconheceu seu empenho e passou a auxiliá-lo financeiramente na aquisição dos títulos. Com apenas 10 títulos fundou, (...) em 1936, um museu-teatro destinado a guardar, restaurar e exhibir filmes: a Cinemateca Francesa (FIGURA 3), que hoje dispõe de mais de 60 mil títulos⁴⁴.

³⁸ VELOSO, Geraldo. 2016, p. 52.

³⁹ LA CARRETA, Marcelo. 2005, p. 25.

⁴⁰ LA CARRETA, Marcelo. 2005, p. 25.

⁴¹ LA CARRETA, Marcelo. 2005, p. 30.

⁴² LA CARRETA, Marcelo. 2005, p. 31.

⁴³ LA CARRETA, Marcelo. 2005, p. 31 apud EISNER, Lotte. p. 175. Original não encontrado pela autora.

⁴⁴ LA CARRETA, Marcelo. 2005, p. 31. Os dados atualizados podem ser encontrados no site: <http://www.cinematheque.fr/> Ativo em 05 jan. 2017.



FIGURA 3 – Fachada da Cinemateca Francesa, em julho de 2016.
 FONTE: Acervo da autora.

O empreendimento de Henri Langlois foi fundamental para a preservação dos filmes e, desde então, foi copiado em todo o mundo. Afinal, além de armazenar os filmes, ele promovia exposições constantes, baseado na premissa de que “um acervo não serve apenas para proteger os filmes do desaparecimento, mas também para divulgá-los, e assim propagar a memória deles”⁴⁵.

A organização maior da classe seria criada [também] por iniciativa de Henri Langlois em 1938: a FIAF (Federação Internacional de Arquivos de Filmes). Com sede em Bruxelas, Bélgica, a associação que no ato de sua criação contava com apenas quatro sócios, hoje conta com mais de 120 instituições, numa mostra da crescente importância da preservação do patrimônio das imagens⁴⁶.

⁴⁵ LA CARRETA, Marcelo. 2005, p. 32.

⁴⁶ LA CARRETA, Marcelo. 2005, p. 33.

O site da FIAF – *International Federation of Film Archives* (www.fiafnet.org) é atualmente uma das maiores fontes de informações para arquivistas audiovisuais de todo o mundo. Entre os principais feitos da FIAF, estão: a padronização das técnicas criadas para a conservação e preservação de filmes, o acesso à investigação histórica, a livre troca de saberes entre cinematecas, bem como o treinamento de funcionários especializados para a árdua tarefa de preservação e exibição do patrimônio audiovisual ⁴⁷.

Ray⁴⁸ afirma que o estudo da preservação de arquivos audiovisuais não funciona como um manual didático frio e distante, mas é algo que “pulsa vida e emoção, sentimento”. Marcelo La Carreta, em sua tese⁴⁹, aponta a importância dos cinéfilos e colecionadores na preservação de filmes e suas tecnologias de leitura em bom estado, ainda que se trate de formatos já obsoletos: filmagens caseiras, fotografias, negativos, entre outros. Em sua definição, obra audiovisual “é um importante documento histórico e, quanto mais o homem toma consciência desse fato, maior é sua preocupação em salvaguardar e preservar esse material”.

A história do audiovisual, apesar de relativamente nova se comparada à de outras formas de arte, revela-se cada vez mais falha, sobretudo quando se refere à contemporaneidade. Tendo o cinema como um importante registro de várias épocas desde o final do século XIX, uma das chaves que podem levar ao autoconhecimento, é salutar que ele alcance as próximas gerações. Devido à sua efemeridade, a salvaguarda de produções audiovisuais é um dos maiores desafios a serem enfrentados⁵⁰.

La Carreta destaca ainda que a importância desse ato de preservar independe de “época, gênero ou formato, seja um filme, um comercial, uma animação, um telejornal, uma gravação de festa de família ou um videogame”. Para ele o produto audiovisual deve ser visto sempre como um documento histórico a ser conservado, uma vez que “se trata de um dos mais fiéis retratos culturais do mundo em que vivemos”⁵¹.

⁴⁷ LA CARRETA, Marcelo. 2005, p. 33.

⁴⁸ Ray, EDMONDSON. 2013, p. 9.

⁴⁹ LA CARRETTA, Marcelo. 2012, p.21.

⁵⁰ LA CARRETTA, Marcelo. 2012, p.11

⁵¹ LA CARRETTA, Marcelo. 2012, p.21

Com tantas vertentes na contemporaneidade, (...) faz-se urgente pensar nos novos desafios que o século enfrentará para manter vivos os documentos audiovisuais. O novo mundo de telas é extremamente complexo, e se o virtual é ideal para propagar, ainda não se mostrou seguro na conservação da obra audiovisual, se é que existe uma solução definitiva para a mais efêmera das artes⁵².

Sabemos que “os documentos audiovisuais são muitas vezes efêmeros por sua própria natureza”⁵³. Os suportes não são facilmente adaptáveis “nos pressupostos de trabalho das bibliotecas, arquivos e museus”⁵⁴ e por isso tamanha foi a dificuldade inicial em se construir um novo tipo de instituição que os abrigasse. O depoimento abaixo, de Rod Wallace — da National Library da Austrália — evoca a década de 50, não tão distante de nós, e como eram tratados os arquivos e as pessoas que deles cuidavam:

As atitudes públicas para com os materiais históricos eram então muito diferentes, em especial no mundo cinematográfico. Nós éramos olhados como doidos, e em muitas ocasiões isso nos era dito. Nunca vou esquecer uma vez que uma sala de projeção cheia de gente da indústria cinematográfica assistiu um programa de filmes antigos recuperados pela Biblioteca e um homem me disse que nós deveríamos ter jogado tudo fora. E as outras pessoas concordaram com ele!⁵⁵

Os arquivos fílmicos, como pudemos perceber, foram se alastrando em forma de locais bem distintos das tradicionais instituições de preservação. Surgindo primeiro na Europa e América do Norte, em meados dos anos 1930. Devagar, em fases, e diante de barreiras que pareciam intransponíveis, o valor do patrimônio audiovisual foi legitimado e teve boa aceitação⁵⁶.

Evidente que, mesmo hoje, nem todos concordarão com a importância da preservação. Mas, a essa altura, muitas outras instituições de guarda dos acervos audiovisuais foram criadas no mundo e sua relevância foi, finalmente, comprovada.

⁵² LA CARRETA, Marcelo. 2012, p.34.

⁵³ EDMONDSON, Ray. 2017, p. 09.

⁵⁴ EDMONDSON, Ray. 2013, p. 93.

⁵⁵ EDMONDSON, Ray. 2013, p. 94.

⁵⁶ EDMONDSON, Ray. 2013, p. 94.

1.2 Contexto brasileiro

Vivemos num país onde mais de noventa por cento da produção realizada no período do cinema silencioso se perdeu, onde o primeiro programa de restauro de filmes da maior cinemateca foi implementado quase cem anos depois da invenção do cinema, onde existem períodos de longa ausência de políticas efetivas de conservação da memória audiovisual⁵⁷.

A frase de Andréia França introduz a realidade crítica que se encontra impregnada nas instituições brasileiras de preservação. Se, internacionalmente, o processo de afirmação do valor dos filmes enquanto patrimônio já foi complexo, não somos sequer capazes de mensurar o que os pioneiros brasileiros passaram.

O cinema no Brasil, assim como no mundo, surgiu com a preocupação de apenas produção de entretenimento — sem um planejamento posterior. La Carreta nos relembra que “esta prática leva nosso cinema a um eterno recomeço, já que não se formam as bases históricas para o futuro aprendido”⁵⁸. Segundo ele, a memória do cinema nacional se sustentou, principalmente, por iniciativas escassas e particulares, que contava com acervos pessoais e também relíquias de família⁵⁹.

O Brasil se interessa pouco pelo seu passado. Essa atitude [...] exprime a vontade de escapar a uma maldição de atraso e miséria. O descaso pelo que existiu explica não só o abandono em que se encontram os arquivos nacionais, mas até a impossibilidade de se criar uma cinemateca. Essa situação dificulta o trabalho do historiador, particularmente o que se dedica a causas sem importância como o cinema brasileiro⁶⁰.

A contestação se baseia no contexto nacional da época, quando Paulo Emílio Salles Gomes se esforça em criar a Fimoteca do Estado de São Paulo, semente do que veio a ser a Cinemateca Brasileira — inspirada no modelo da *Cinémathèque Française* e uma das únicas ainda hoje amparadas pelo governo nacional⁶¹.

⁵⁷ FRANÇA, Andréia. 2016, p.67

⁵⁸ LA CARRETA, Marcelo. 2005, p. 77.

⁵⁹ LA CARRETA, Marcelo. 2005, p. 77.

⁶⁰ GOMES, Paulo Emílio Salles. 1980, p. 22.

⁶¹ LA CARRETA, Marcelo. 2005, p. 82.

Paulo Emílio garimpou filmes nacionais por todo o país e seu plano era fundar um arquivo de guarda do cinema brasileiro que funcionasse de forma ininterrupta, o que foi cumprido mesmo com o incêndio que ocorreu na instituição em 1957, pelo qual ele e todos lamentaram⁶². Pode-se imaginar que a inauguração dessa grande instituição tenha mudado a consciência sobre a importância de preservação dos filmes nacionais, mas não foi bem assim.

A Cinemateca Brasileira representou, certamente, um grande marco para o país. Ainda hoje, a instituição é referência em "preservação e difusão do patrimônio cinematográfico e audiovisual em todas as suas manifestações, [e] tem tido a satisfação de desenvolver diversos projetos de recuperação, valorização e acesso de acervos variados"⁶³.

Esses mais de 70 anos de atividade [da Cinemateca Brasileira] não transcorreram sem problemas, que refletem tanto a questão mais geral da necessidade de se valorizar a memória e a cultura nacionais quanto as dificuldades de se estabelecer padrões rigorosos para o cumprimento de uma audaciosa missão em meio à instabilidade financeira e política.⁶⁴

Os desafios, então, não fazem parte, apenas, das instituições não subsidiadas pelo governo — eles são dependentes do processo de preservação desde seu surgimento e importante avanço. Atualmente, a Cinemateca Brasileira possui um acervo significativo: com aproximadamente 245 mil rolos de filmes, correspondentes a 30 mil títulos. "São obras de ficção, documentários, (...) filmes publicitários e registros familiares, nacionais e estrangeiros, produzidos desde 1910. [Eles] foram incorporados à Cinemateca através de depósito voluntário, doação e depósito legal"⁶⁵ possuindo, inclusive, acervos mineiros da obra de Humberto Mauro pertencentes à Escola de Belas Artes da UFMG e depositados para maior segurança e cuidado com os filmes.

Desde 2003 acoplada à Secretaria do Audiovisual, a Cinemateca Brasileira deu, nos anos entre 2008 e 2013, "um grande impulso nas suas atividades de preservação e

⁶² LA CARRETA, Marcelo. 2005, p. 81 e 83.

⁶³ BRASILEIRA, Cinemateca. Em prefácio à edição traduzida do livro *O dilema digital*, sem página definida. Atualização dos anos feita pela autora, baseada nas informações do site: <http://www.cinemateca.gov.br>. Ativo em: 09 jan. 2017.

⁶⁴ IDEM.

⁶⁵ <http://www.cinemateca.gov.br/pagina/preservacao-audiovisual>. Ativo em: 09 jan. 2017.

difusão de acervos, bem como na infraestrutura"⁶⁶, graças aos recursos investidos pelo próprio Ministério da Cultura.

Cabe ressaltar que, mesmo ela gozando de auxílio do governo federal brasileiro para seu funcionamento e manutenção, após esse frutífero período, a Cinemateca Brasileira passou por uma crise "cujos desdobramentos levaram à drástica redução das equipes e, por conseguinte, da sua capacidade operacional"⁶⁷.

Um evento significativo — e que até mesmo impulsionou este trabalho — aconteceu no início do ano de 2016: um incêndio que acometeu um dos galpões da Cinemateca Brasileira (FIGURA 4), "resultando na perda de 1.003 rolos de filmes em nitrato de celulose, referentes a 731 títulos"⁶⁸. Este foi o quarto, desde sua criação e segundo a instituição, desde esse acontecimento, "os poucos recursos disponíveis são concentrados na duplicação emergencial do acervo audiovisual e normalização mínima das rotinas de cada setor"⁶⁹.



FIGURA 5 - Reportagem sobre o incêndio em 4 de fevereiro de 2016.

Fonte: <http://cinemacomrapadura.com.br/noticias/401882/1-000-rolos-de-filmes-foram-queimados-em-incendio-na-cinemateca-brasileira/> Acesso em: 29 abril 2017.

⁶⁶ <http://www.cinemateca.gov.br/pagina/a-cinemateca-historia>. Ativo em: 09 jan. 2017.

⁶⁷ <http://www.cinemateca.gov.br/pagina/a-cinemateca-historia>. Ativo em: 09 jan. 2017.

⁶⁸ <http://www.cinemateca.gov.br/pagina/a-cinemateca-historia>. Ativo em: 09 jan. 2017.

⁶⁹ <http://www.cinemateca.gov.br/pagina/a-cinemateca-historia>. Ativo em: 09 jan. 2017.

Ora, se estamos defendendo que o cinema deve ser preservado por intermédio das instituições de guarda, é responsabilidade delas garantir a integridade dos arquivos, mesmo diante da impossibilidade de impedir que acidentes aconteçam. Se não somos capazes de evitar uma tragédia, devemos repensar nosso próprio campo de atuação e o potencial de nossas ações. Sendo assim, seria interessante pensar que, se existisse algum tipo de cópia digital desses filmes, a perda do suporte não seria tão drástica. Sabemos que a experiência não é a mesma, mas, afinal, preserva-se mais e principalmente o *conteúdo* presente na materialidade da película. Isso será melhor explicado no capítulo a seguir.

Já a Cinemateca do MAM, no Rio de Janeiro, completou 60 anos em 2015. Foi criada para acolher o cinema nas exposições do museu de arte moderna, e, ao longo dos anos, "estruturou-se em arquivo de filmes, centro de documentação correlata e sala de exibição audiovisual dos diferentes suportes, formatos, bitolas e padrões que o cinema e as demais formas de imagens em movimento assumiram ao longo do tempo, cobrindo película, magnético e digital"⁷⁰.

Seu acervo atual e acumulado é de aproximadamente "80 mil rolos de filmes, 60 mil títulos em mídias magnéticas e digitais, (...) além de catálogos, filmografias, obras de referência, brinquedos, equipamentos e diversas outras tipologias documentais"⁷¹.

Outras instituições menores, ligadas a escolas e empresas, também foram criadas regionalmente em todo o país na década de 80⁷².

Porém, nos anos 2000 ocorre a desestruturação da Cinemateca do MAM (...), uma das mais antigas do país. Passando por várias dificuldades, acabou desmontada, pois o conselho curador do MAM quis desfazer-se do acervo. A Cinemateca Brasileira por sua vez passou por dificuldades semelhantes, que quase a levaram à falência em 1984, mostrando que esta mentalidade conservacionista não está tão consolidada assim⁷³.

⁷⁰ http://mamrio.org.br/wp/museu_cinemateca/apresentacao. Ativo em: 09 jan. 2017.

⁷¹ http://mamrio.org.br/wp/museu_cinemateca/apresentacao. Ativo em: 09 jan. 2017.

⁷² LA CARRETA, Marcelo. 2005, p. 83.

⁷³ LA CARRETA, Marcelo. 2005, p. 83.

Laura Bezerra, Professora e ex-presidente da ABPA – Associação Brasileira de Preservação Audiovisual, declara que os anos 2000 marcaram "o início de um processo de amadurecimento da preservação audiovisual no Brasil"⁷⁴. Apesar de lento, tal processo ganhou força principalmente em meio aos produtores e realizadores que passaram a perceber a importância da preservação na cadeia produtiva do cinema⁷⁵.

A própria criação da ABPA significou grande avanço, visto que, até então, as iniciativas na área da preservação dos filmes eram feitas de forma isolada pelas instituições. Trata-se de uma junção dos "órgãos governamentais, instituições, associações de classe, técnicos, pesquisadores, historiadores, colecionadores, cineastas, jornalistas e os participantes da CINEOP – Mostra de Cinema de Ouro Preto"⁷⁶ em benefício da preservação do cinema nacional por meio de estudos e troca de informações.

Até então eram raras as universidades de cinema e audiovisual no país que possuíam em sua grade curricular disciplinas de formação nas áreas de conservação e preservação. Isso representa uma decisão que desrespeita a própria existência do curso — pensando que ele jamais poderia ser ofertado se filmes não tivessem sido preservados outrora. Esse cenário tem mudado com a resolução número 10/2006 do Conselho Nacional de Educação ⁷⁷, que propôs a **obrigatoriedade** de inserção da preservação audiovisual como disciplina obrigatória nos cursos de cinema do país⁷⁸.

Desde então, diversos trabalhos acadêmicos foram escritos sobre o tema e elevou-se o nível de reflexão sobre a preservação para além do espaço das instituições detentoras de acervos de imagens em movimento. Aumentou também a percepção da preservação audiovisual como área de atuação profissional (...)⁷⁹.

⁷⁴ BEZERRA, Laura. 2016, p. 28.

⁷⁵ BEZERRA, Laura. 2016, p. 28.

⁷⁶ <https://abpablog.wordpress.com/>. Ativo em: 09 jan. 2017.

⁷⁷ Que pode ser encontrada em: <http://portal.mec.gov.br/busca-geral/323-secretarias-112877938/orgaos-vinculados-82187207/12709-resolucoes-ces-2006> ou http://portal.mec.gov.br/cne/arquivos/pdf/rces10_06.pdf. Acesso em: 20 nov 2017.

⁷⁸ BEZERRA, Laura. 2016, p. 28.

⁷⁹ BEZERRA, Laura. 2016, p. 28.

O trecho é próspero e aponta para uma mudança na mentalidade que parece reconhecer a importância da preservação do patrimônio cinematográfico nacional. Porém, cabe lembrar que o Brasil possui condições adversas à preservação audiovisual. O clima quente e tropical exige cuidados maiores na criação dos ambientes de reserva técnica.

A cinemateca do MAM (FIGURA 6), por exemplo, está instalada no Aterro do Flamengo, uma área instável a poucos metros do mar — o que aumenta a salinidade e umidade do museu.

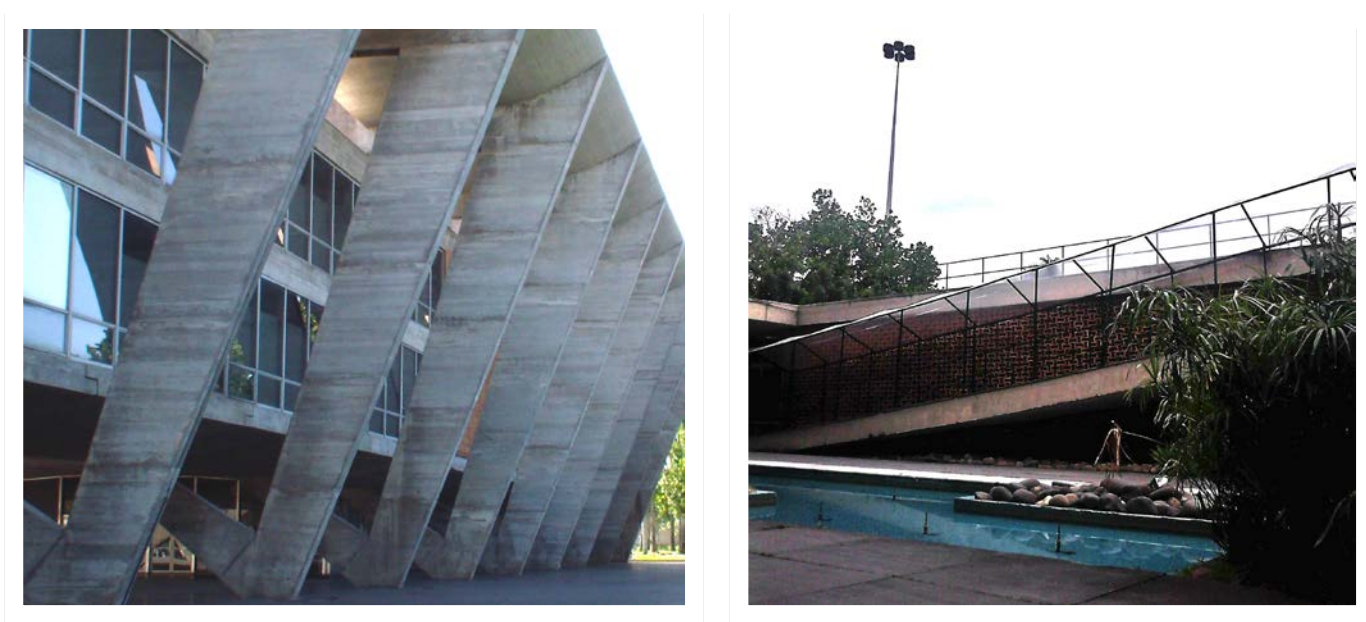


FIGURA 6 – Fachada do MAM e prédio da Cinemateca em setembro de 2016.
FONTE: Acervo da autora.



FIGURA 7 – Reserva técnica da Cinemateca do MAM em julho de 2013.
FONTE: Acervo da autora.

Porém, esses problemas são ínfimos quando se observa a carência do país no principal requisito para que a preservação dos filmes, efetivamente, aconteça: a criação de políticas públicas eficazes, bem como o investimento financeiro (inicial e de manutenção).

É o que diz Solange Stecz no livro "Reflexões sobre a preservação audiovisual": "A salvaguarda e a preservação do patrimônio audiovisual carecem de políticas públicas que considerem sua importância como elemento essencial e estratégico do desenvolvimento da sociedade e da cultura brasileira"⁸⁰. Filmes tantos já se perderam, "tanto pelo fim de sua vida comercial quanto pela fragilidade de seus suportes e pela falta de conscientização sobre sua importância para a memória e história do país"⁸¹.

Hoje a Cinemateca do MAM sobrevive, também, pelo amor que os funcionários e voluntários dedicam ao trabalho com os filmes. Às vezes investem do próprio bolso na aquisição de insumos, procuram incessantemente por programas de financiamento, estudam as inovações muitas da área, e participam de eventos em que a troca de saberes entre instituições pode colaborar para o propósito único da perenização dos filmes.

José Quental, ex-funcionário da cinemateca do MAM do estado do Rio de Janeiro, afirma que a realidade se funda em "reiteradamente barganhar recursos mínimos para realizar parte das atividades mais corriqueiras de um arquivo de imagens em movimento, [com isso ele percebe que] a ideia do audiovisual enquanto patrimônio não é assim tão consolidada"⁸².

Se, por um lado, a afirmação das imagens em movimento como parte do patrimônio cultural que precisa ser conservado e transmitido para as gerações futuras, não é nenhuma novidade - basta lembrar que o Brasil é signatário da Recomendação pela Salvaguarda e Conservação das Imagens em Movimento, da UNESCO, de 1908 -, por outro lado as políticas culturais têm sido historicamente insuficientes para cumprir com essa missão⁸³.

⁸⁰ STECZ, Solange Straube. 2016, p.31.

⁸¹ STECZ, Solange Straube. 2016, p.31.

⁸² QUENTAL, José. 2016, p.35.

⁸³ QUENTAL, José. 2016, p.35.

O exemplo da Cinemateca do MAM é apenas um, dentre diversos. No MIS BH, antigo Crav, encontram-se casos semelhantes. Os funcionários, normalmente concursados da prefeitura da cidade de Belo Horizonte, também trabalham intensamente na manutenção do acervo.

Durante o andamento dessa pesquisa, foi possível acompanhar, inclusive, uma ação contra o MIS-BH. A sede situada em um casarão histórico totalmente adaptado para as reservas técnicas recebeu ordens da prefeitura da cidade para ser transferida, sem aviso, para um local ainda sem condições de ambientar o acervo e a um custo altíssimo — que sempre dizem não ter para investir em preservação (FIGURA 7). O andamento desse processo foi finalizado, após vários cineclubes e associações locais lutarem e reivindicarem a estadia do acervo no local apropriado, bem como entrada de alguns vereadores da cidade em defesa da causa (FIGURA 8) que foi encerrada em setembro do mesmo ano e a sede foi mantida no casarão histórico da Avenida Álvares Cabral, já adaptado para o acervo.



FIGURA 7 – Reportagem sobre mudança do MIS

Fonte: <http://observatoriodiversidade.org.br/site/prefeitura-de-belo-horizonte-determina-mudanca-de-endereco-do-museu-da-imagem-e-do-som/> Acesso em: 13 fev 2018.

Mudança de endereço do Museu da Imagem e do Som é criticada por vereadores

Os parlamentares questionam a incapacidade do Cine Santa Tereza em receber o acervo do MIS

por **Encontro Digital** 09/05/2017 12:15



FIGURA 8 – Reportagem sobre mudança do MIS

Fonte: <https://www.revistaencontro.com.br/canal/atualidades/2017/05/mudanca-de-endereco-do-museu-da-imagem-e-do-som-e-criticada-por-veread.html> Acesso em: 13 fev 2018

No que se refere às técnicas de preservação do audiovisual analógico, o conhecimento produzido até hoje é bastante vasto. Valendo-se dos saberes advindos da preservação de arte, do trabalho com polímeros e das condições ambientais da teoria clássica de preservação, muito já se avançou em termos de técnicas e confiabilidade. Hernani Heffner, o curador e conservador-chefe da Cinemateca do MAM, treina voluntários e interessados há muitos anos. No MIS em BH vários cursos são ministrados ao longo do ano e é um conhecimento já bastante vasto em bibliotecas digitais das próprias universidades.

Mais do que isso, já se sabe também das razões pelas quais os filmes são dignos de serem preservados:

qualquer filme é importante sob a ótica de uma análise histórica, pois todos registram não só imagens, mas ideias, costumes, gestos da humanidade daquele período. Porém, o cinema revelou-se, com o passar dos anos, a mais efêmera das artes: traiçoeira, enganou os que acreditavam em sua eternidade, sendo sua mídia um dos mais precários suportes já inventados. Quando se percebeu este fato, restavam apenas 10% das imagens feitas pelos pioneiros. Desde então existe a corrida contra o fim dos filmes para se salvar o que se puder do maior inimigo da memória audiovisual: o tempo⁸⁴.

Dessa forma, sejamos práticos. Muito já foi dito sobre o fato quase óbvio de que é necessário preservar, mas, num país como o Brasil é totalmente improvável que se preserve tudo.

Nossas cinematecas e instituições simplesmente não teriam espaço, pessoal, orçamento para todos os longas, todos os curtas, todos os cinejornais, todos de tudo. (...) Porque afinal quem escolhe o que se preservará, o que se restaurará (ainda mais grave)? Sob quais critérios? Obras relevantes? Para quem? Geralmente é sob este tipo de pressão que a coisa começa a ficar estranha, porque passamos por conceitos voláteis como o de relevância, importância, permanência, que significam uma coisa para cada pessoa.⁸⁵

Em vários países, já não se faz tão necessário argumentar que os filmes e documentos audiovisuais *precisam* ser preservados e necessitam de estocagem correta e inúmeros cuidados em sua salvaguarda que lhe conferem a sobrevivência⁸⁶. Mas La Carreta nos adverte que estamos nos tornando uma cultura incapaz de preservar as próprias imagens, que significam parte relevante da história de cada geração⁸⁷.

Riquíssima de elementos, a memória audiovisual é fundamental para compor o novo mundo de imagens em que vivemos. A preservação do cinema está submetida à política, mas acima desta está a real função das imagens, imprescindíveis para a compreensão do passado, base para o futuro imagético e digital⁸⁸.

⁸⁴ LA CARRETA, Marcelo. 2005, p. 7.

⁸⁵ LA CARRETA, Marcelo. 2005, p18.

⁸⁶ EDMONDSON, Ray. 2013, p. 37

⁸⁷ LA CARRETA, Marcelo. 2005, p18

⁸⁸ LA CARRETA, Marcelo. 2005, p18

A preservação de arquivos audiovisuais vem sendo feita por diferentes tipos de instituição e se expande lentamente. As próprias produtoras já têm maior consciência de que se faz necessário salvaguardar seus próprios arquivos — acatando o argumento de que a rememoração é fator importante no processo da preservação — considerando-se a perspectiva de que a história não pode ser esquecida e as perdas devem sempre ser consideradas.

1.3 Relevância da memória

É habitual encontrarmos o termo memória sendo utilizado em contextos distintos e vemos uma recente preocupação com o seu resgate, além de um constante medo de sua perda. Vemos até mesmo os mais jovens tomando remédios e se prevenindo desse mal, mas sequer conseguem se lembrar do próprio número de telefone sem a ajuda de algum dispositivo que guarde informações.

Ao discorrer sobre o tema, Marcos Palacios⁸⁹ afirmou que falar de memória “está na moda”, principalmente nos cursos de comunicação. Sabemos que a internet não nos deixa esquecer, uma vez que produzimos conteúdo quase o tempo todo, nas mais diversas plataformas *online*. Giselle Beiguelman⁹⁰ nos lembra que tais serviços podem simplesmente desaparecer a qualquer momento, já tivemos essa experiência com o *Orkut* — uma rede social que se tornou obsoleta e, ao ser eliminada, levou consigo fotos e dados dos usuários.

A obsolescência dos equipamentos também é rápida e “uma estranha nostalgia de um passado não vivido” ocupa até mesmo os meios de consumo. É momento oportuno para pensarmos nessas memórias “instáveis” que se findam juntamente à duração dos equipamentos e no tipo de memória que estamos construindo na *web*, “onde o presente imediato parece ser o tempo essencial”⁹¹.

⁸⁹ PALACIOS, Marcos Silva. Palestra na abertura do 10º ALCAR – Encontro Nacional de História da Mídia, com o tema “A memória na era Digital”. 03 jun 2015.

⁹⁰ BEIGUELMAN, Giselle. 2014, p 12.

⁹¹ BEIGUELMAN, Giselle. 2014, p 12.

Já no cinema, tal relação é ainda mais antiga: há quem diga⁹² que “o cinema é memória e o é de muitas maneiras”, sendo a sétima arte tratada como memória audiovisual do mundo. Podemos perceber isso não somente nos filmes documentais, mas a própria ficção se apropria de narrativas da memória ao fazer *flashbacks* e rememorar algo da cena do vivido que ficou no passado.

Nunca se falou tanto em memória como hoje em dia, e nunca foi tão difícil ter acesso ao nosso passado recente. (...) Poucas palavras tornaram-se tão corriqueiras (...) como 'memória'. Até bem pouco tempo confinada aos campos da reflexão historiográfica, neurológica e psicanalítica, a memória converteu-se num aspecto elementar do cotidiano. Tornou-se uma espécie de dado quantificável, uma medida e até um indicador do *status* social de alguém. Existe um fetiche da *memória* como *coisa*: quanto de memória tem seu computador? E sua câmera? E seu celular? Tudo isso? Só isso?... Compram-se memórias, transferem-se memórias, apagam-se e perdem-se memórias⁹³.

Percebemos então que é, talvez, impossível definir o texto de memória e garantir-lhe uma especificidade sem questões, sem hibridismos, sem contaminações com outros discursos. Sendo assim, cineastas também possuem esse “poder” de reconstrução e o trabalho da memória se torna uma operação transformadora.

Outras definições apresentam o arquivo como um “lugar físico que abriga o destino dessa espécie de rastro que cuidadosamente distinguimos do rastro cerebral e do rastro afetivo, a saber, o rastro documental”⁹⁴.

o arquivo não deve ser visto apenas como depósito, lugar onde se deixam objetos inúteis, remanejados para os porões do tecido social, instância onde se guardam enunciados sem vida, acumulados de forma mecânica (...) O arquivo só se dará a conhecer, após sua elaboração, a leitores de um outro tempo. Os enunciados serão continuamente ressignificados⁹⁵.

Podem ser chamados “arquivos” todo tipo de material, temos, portanto, que a chamada técnica arquivística não determina o momento do registro, mas a instituição mesma do acontecimento arquivável.

⁹² ÁLVAREZ, Efrén Cuevas.

⁹³ BEIGUELMAN, Giselle. 2014, p 13.

⁹⁴ RICOEUR, Paul. 2007, p.177.

⁹⁵ MENEZES, Roniere. 2014, p. 235.

No caso dos arquivos audiovisuais, é importante ressaltar que ainda não há uma palavra final sobre o assunto, lembrando que esse tipo de arquivo é muitas vezes efêmero por sua própria natureza. Chamaremos, portanto, de audiovisual todas as imagens em movimento acompanhadas por qualquer tipo de áudio. Intitularemos, também, *suporte* os itens físicos (VHS, CD, DVD, HD) e *mídia* o conteúdo da informação específico (o arquivo digital, o vídeo em si, *codecs* e afins). Ambos podem estar diretamente unidos, mas se faz necessário que tenhamos o acesso aos dois. O digital, como veremos adiante, complica tais definições por sua falta de materialidade.

O cinema sempre teve sua sobrevivência associada à superação de obstáculos tecnológicos que, ao mesmo tempo em que permitiam sua evolução, jogavam ao esquecimento tudo o que já fora feito para dar lugar a novidades. O cinema é totalmente dependente de seu suporte, e como observou o crítico e historiador de cinema Leonard Maltin, a ironia é que, apesar de mais novo, o filme é mais frágil do que qualquer outra mídia.⁹⁶

⁹⁶ Documentário *Salvadores de imagens (Keepers of the frame, 1999)*, da WinStar Productions.



CAPÍTULO 2

PRECISAMOS FALAR SOBRE O DIGITAL

Precisamos falar sobre o digital porque temos observado um avanço rápido das tecnologias, porque milhares de *terabytes* de informação são produzidos todos os dias, porque as taxas de *uploads* de fotos e vídeos em redes sociais nunca foram tão altas, porque *pendrives* e HDs externos queimam. Precisamos falar sobre o digital porque arquivos são corrompidos sem explicação aparente, porque incêndios acontecem fora do nosso alcance e nada sabemos fazer, porque os hábitos de consumo do espectador mudaram e continuam em transição, porque pessoas assistem filmes em celulares e *tablets*. Precisamos falar sobre o digital porque sequer temos demanda por essa busca de termo no Google.

As imagens (FIGURAS 9, 10, 11 e 12) a seguir foram salvas em abril de 2017, durante o desenvolvimento desta dissertação. A busca foi feita para que pudéssemos metrificar a grande maioria do conteúdo produzido em relação a esse assunto, bem como acessá-los, mas tamanha foi a surpresa ao ver que o termo vem sendo cada vez menos pesquisado se considerarmos os últimos 13 anos (tempo limite que o Google oferece). E isso levando em conta que há 10 anos o acesso à internet não era tão democrático quanto hoje, principalmente no Brasil, como podemos observar também a seguir.



FIGURA 9 - Infográfico - dados sobre o acesso de internet no Brasil.

Fonte: http://blog.pmweb.com.br/wp-content/uploads/infografico_internet_no_brasil.png

Acesso em: 29 abr. 17.

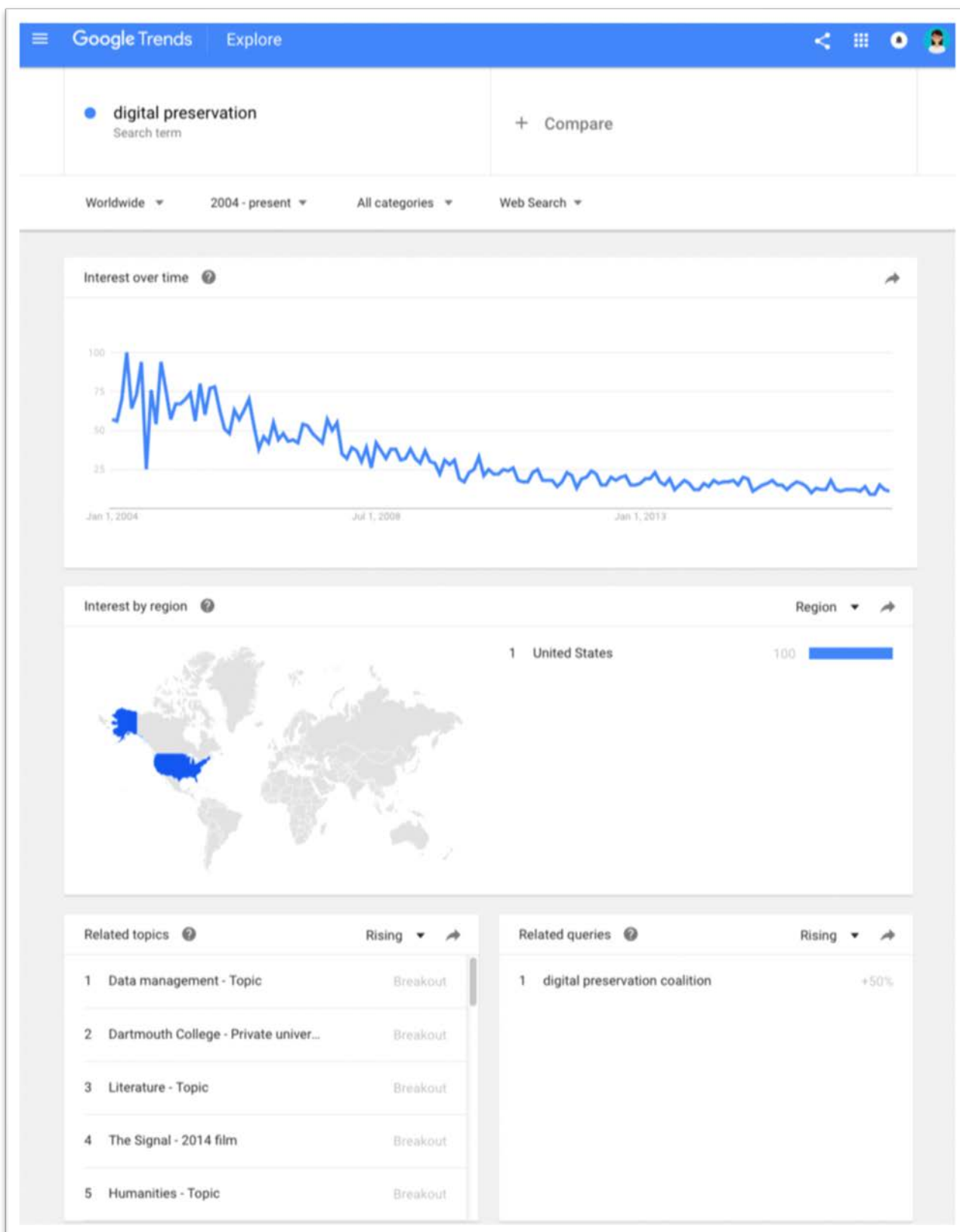


FIGURA 10 - Busca de termos no Google Trends.
Fonte: trends.google.com/. Acesso em: 29 abr 17.

O gráfico acima (FIGURA 10) foi gerado pelo *Google Trends*, uma ferramenta de busca de palavras chave e termos mais procurados no servidor de busca da Google — o mais utilizado em todo o mundo. Inicialmente foi feita a busca pelos termos "preservação digital" em inglês "*digital preservation*", desde 2004 e em todo o mundo. Ele também apresenta o local de origem das pesquisas e palavras relacionadas a essas buscas. Por aí já podemos perceber que não se trata de um assunto muito procurado, e sua maioria está localizada em apenas um país. Quando a pesquisa é feita utilizando apenas a expressão em Língua Portuguesa, com público brasileiro, os resultados são inexistentes e o gráfico sequer tem informações para exibir em virtude da baixa procura.

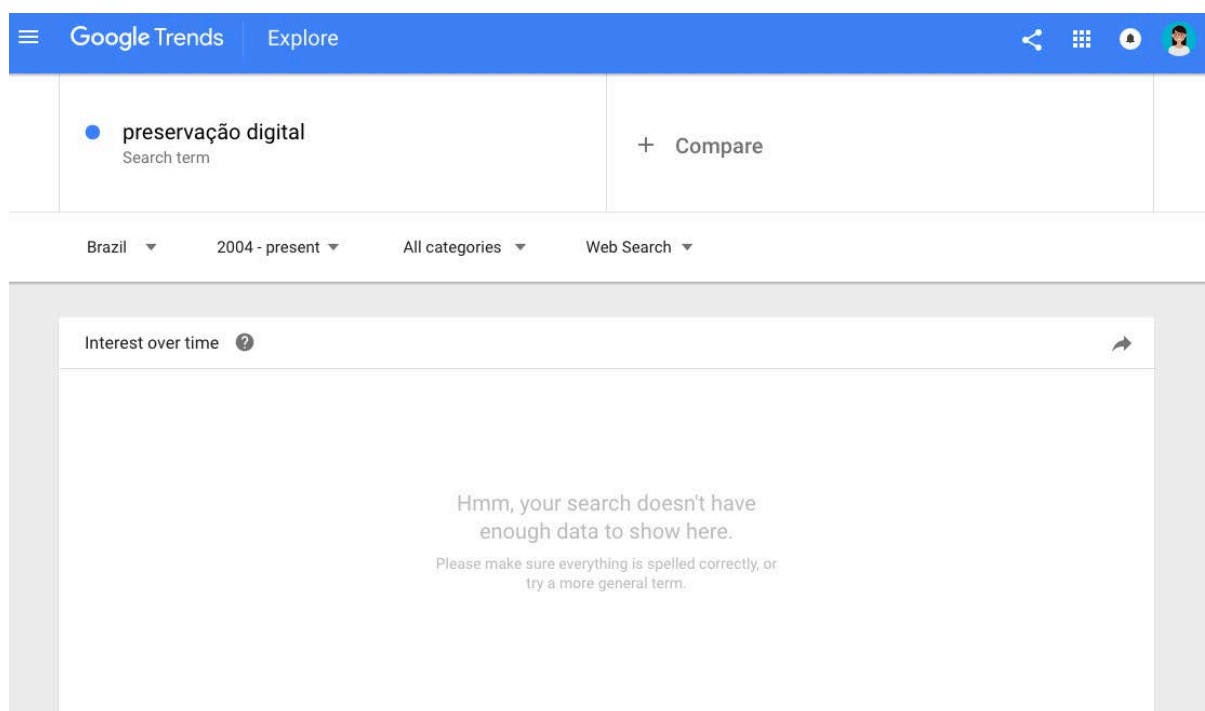


FIGURA 11 - Falha na busca de termos no Google Trends.
Fonte: trends.google.com/. Acesso em: 29 abr 17.

Mais assustador ainda é perceber que a mesma situação acontece quando a pesquisa vai de encontro ao tema deste trabalho, na busca dos termos "preservação audiovisual" que não possuem relevância de pesquisa nem mesmo em Língua Inglesa. Porque "preservação digital" ainda é termo extremamente amplo e engloba todo e qualquer tipo de arquivo — textos, imagens, vídeos, planilhas e tudo mais.

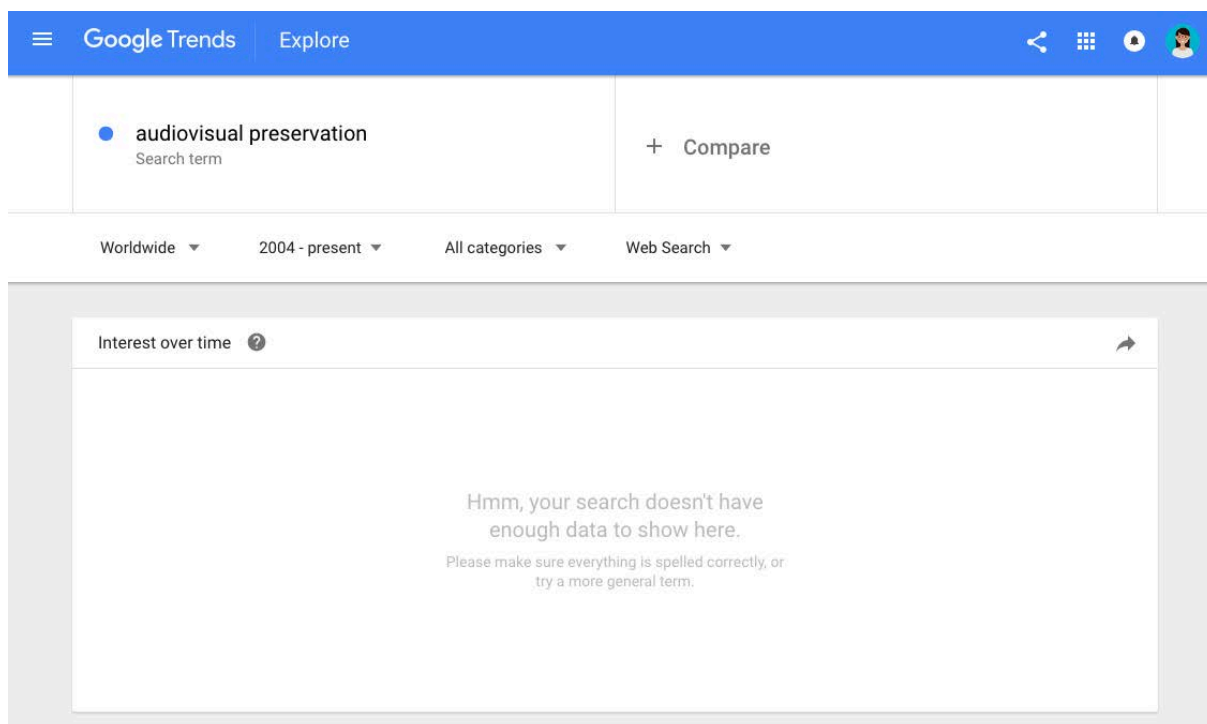


FIGURA 12 - Falha na busca de termos no Google Trends.
 Fonte: trends.google.com/. Acesso em: 29 abr 17.

Tais buscas não significam que ninguém se importa com o assunto, mas esses dados significam muito no discurso que estamos propondo de que se faz necessário preocupar com a situação de nossos acervos — principalmente os audiovisuais que agora enfrentam o desafio da transição para o digital.

Após a compreensão do primeiro capítulo e também dessa pequena introdução dos termos, dessa forma, agora já podemos considerar objetivo maior deste trabalho: colaborar com o pensamento de que as mídias atuais que usamos com muita frequência são altamente instáveis e precisam ser constantemente atualizadas.

No Brasil, o Cine OP⁹⁷ de 2017 abordou o tema das Emergências Digitais. O evento reuniu especialistas de diversas áreas correspondentes e promoveu discussões importantíssimas principalmente nos contextos brasileiro e latino americano.

⁹⁷ Festival de Cinema da cidade de Ouro Preto, Minas Gerais. Realizado anualmente no mês de junho ou julho e sempre voltado à temática da preservação audiovisual, trazendo nomes relevantes para debates.

O Cine OP é uma mostra de cinema que já possui mais de 10 anos de existência e sempre trabalha em suas mesas e discussões as questões ligadas à preservação de acervos audiovisuais. Os assuntos ligados à preservação e debates que ocorreram nos anos anteriores podem ser encontrados com facilidade no site⁹⁸.

O digital, visto em seus primórdios como solução de alguns problemas, acabou se mostrando ainda mais problemático que o analógico. Ainda temos muito o que discutir em termos de confiabilidade e armazenamento desses arquivos. Sem falar na obsolescência programada, assunto recorrente da modernidade que nos mostra que as questões de mercado ainda estão acima das questões de preservação.

2.1 A estrutura do arquivo digital

Atualmente, grande parte (ou quase a totalidade) da produção intelectual é feita por meio de ferramentas digitais — computadores, *tablets*, celulares e afins, com *softwares* e aplicativos específicos e principalmente por meio do acesso à *web*. A facilidade de criação e disseminação desses arquivos é um fator relevante quando se decide utilizar tais ferramentas.⁹⁹

No entanto, o material digital carrega consigo um problema estrutural que coloca em risco a sua longevidade. Embora um documento digital possa (em tese) ser copiado infinitas vezes sem qualquer perda de qualidade, este exige a presença de um contexto tecnológico para que possa ser consumido de forma inteligível por um ser humano. Esta dependência tecnológica torna-o vulnerável à rápida obsolescência a que geralmente a tecnologia está sujeita.¹⁰⁰

São muitos os exemplos de obsolescência que podemos citar ao longo dos anos. No caso do vídeo/cinema amador, basta se lembrar das fitas tipo VHS que fizeram parte da infância dos anos 90. Para que seja possível assisti-las ainda hoje, se faz necessário um vídeo leitor de VHS em perfeito funcionamento — o que é raro de encontrar. É necessário, também, que a fita esteja em boas condições de

⁹⁸ Temas anteriores podem ser encontrados aqui: <http://cineop.com.br/a-mostra/edicoes-anteriores>. Acesso em 13 fev 2018.

⁹⁹ FERREIRA, Miguel. 2006, p.17.

¹⁰⁰ FERREIRA, Miguel. 2006, p.17-18.

conservação — sem mofo, arranhões, pedaços arrebatados ou mesmo "mastigados", fato que era comum quando o próprio aparelho "engolia as VHS e acabava amassando a fita magnética".

Com isso, hoje temos como obsoleto o formato de vídeo mais duradouro que já existiu: o VHS. A gravação magnética revolucionou não somente o áudio, mas criou o termo vídeo (do latim *eu vejo*) que se diferenciava do cinema por democratizar os meios de produção e, posteriormente, por sua narrativa, hoje é impossível pensar o audiovisual sem a gravação magnética¹⁰¹.

Os reprodutores de VHS se popularizaram na década de 70, eles eram capazes de reproduzir fitas e gravar os programas de TV. Posteriormente, as filmadoras começaram a ser comercializadas — inicialmente apenas as famílias mais ricas tinham acesso, posteriormente, houve a democratização. Um grande diferencial da gravação em vídeo era o tempo, normalmente uma hora ou duas, enquanto as gravações anteriores em películas duravam, apenas, alguns minutos¹⁰².

Ainda nos anos 90, o VHS já demonstrava sinais de que também não se tratava de uma mídia duradoura em termos de conservação. Com o passar do tempo, as imagens apresentam "chuviscos" principalmente nas bordas, as imagens perdem as cores, a fita "masca" o conteúdo pelo simples fato de ter sido mal colocada no aparelho e seu suporte plástico é um ambiente ideal para propagação de fungos¹⁰³.

Por outro lado, também não podemos defender o digital como um suporte seguro. O DVD também já revela há bastante tempo sinais claros de sua inerente efemeridade. Sua data de validade foi dada como 'indeterminada'. "Indeterminada não por durar muito, mas por ser impossível determinar sua duração que pode ser entre dez anos ou poucos meses, independentemente da qualidade da mídia"¹⁰⁴. Importante ressaltar que essa mídia foi pensada comercialmente no atendimento da nostalgia e

¹⁰¹ LA CARRETA, Marcelo. 2012, p.77.

¹⁰² DIOGO, Lígia. 2010, p. 87.

¹⁰³ LA CARRETA, Marcelo. 2012, p.110.

¹⁰⁴ LA CARRETA, Marcelo. 2012, p.112.

coleccionismo. Trata-se de uma “aposta arriscada criar mediatecas em casa com uma mídia tão vulnerável”¹⁰⁵.

Sendo assim, destaca-se que o problema com os formatos de arquivos que usamos hoje é de mão dupla. Em primeiro lugar, pela própria mídia, se ela vai durar “tempo suficiente para a integridade dos dados e, em segundo, mesmo que a mídia dure até a próxima década ou mais, será necessário o hardware e software para ler os arquivos que estarão disponíveis”¹⁰⁶.

Porém, sabemos que a "obsolescência tecnológica não se manifesta somente ao nível dos suportes físicos".

No domínio digital, todo o tipo de material tem obrigatoriamente de respeitar as regras de um determinado formato. Isto permite que as aplicações de software sejam capazes de abrir e interpretar adequadamente a informação armazenada. À medida que o software vai evoluindo, também os formatos por ele produzidos vão sofrendo alterações. É bastante comum encontrar aplicações de software capazes de carregar os ficheiros produzidos por versões anteriores dessa mesma aplicação. No entanto, essa capacidade raramente vai além das duas versões precedentes.¹⁰⁷

Os dados acima são de 2006, mas ainda hoje permanecem válidos. Uma das grandes alternativas válidas para a preservação do digital, o LTO¹⁰⁸, também possui esse problema: trabalha com gerações, e as mais recentes não leem ou acessam as mais antigas.

E, antes mesmo de saber como funciona a preservação dos filmes binários, se faz necessário compreender a natureza dos arquivos digitais e como eles funcionam, para, então, ser possível propor algo que persista e acompanhe os avanços tecnológicos.

Vamos à definição de qual seria, então, a anatomia de um objeto digital — um arquivo. Miguel cita a definição de K. Thibodeau na qual um ‘objeto digital’ é estabelecido “como todo e qualquer objeto de informação que possa ser

¹⁰⁵ LA CARRETA, Marcelo. 2012, p.112.

¹⁰⁶ LA CARRETA, Marcelo. 2012, p.113.

¹⁰⁷ FERREIRA, Miguel. 2006, p.19.

¹⁰⁸ LTO significa *Linear Tape-Open*, em português, fita aberta linear.

representado por meio de uma sequência de dígitos binários”¹⁰⁹ termo do inglês *bit stream*. O autor ainda considera essa definição como aceitável nos dois principais contextos do digital: os nativos, informações que nasceram digitais, bem como também os digitalizados, a informação procedente de suportes analógicos¹¹⁰.

É comum perder-se nessa definição, uma vez que ela não é suficiente sozinha, deve-se entender o que são os binários e como eles funcionam, para, então, tratar das definições.

Os arquivos se apresentam em formato de textos, planilhas, imagens vetoriais, fotografias, áudios, vídeos, páginas da web, códigos fonte, dentre tantos outros.

Para explicar o funcionamento do digital, retornemos ao analógico. Uma mesma fotografia no papel é diferente de uma fotografia no CD, que, por sua vez, é diferente da fotografia do pendrive e do *e-mail*. O mesmo arquivo inscrito em locais diferentes assume novas características e demanda novos meios de leitura. Isso também acontece com as películas de vídeo: de nitrato, acetato e poliéster, a maneira de ler nesse caso é igual, sob a incidência de luz, mas a forma como uma mesma imagem se comporta nos três suportes é totalmente diferente.

Para que possamos acessar esses arquivos, existe um procedimento correto de leitura que precisa ocorrer. E ele demanda uma máquina (*hardware*) juntamente a um programa (*software*) e com isso tem-se o objeto digital que também é físico. A parte física é o que costumamos chamar de suporte.

A parte física normalmente pode ser lida pelo *hardware* e é ele o responsável por permitir o acesso aos arquivos. Os arquivos, por sua vez, seguem uma estrutura imposta pelo *software*, com regras próprias para que ele possa ser lido e acessado. "A imagem que posteriormente se forma na mente do receptor constitui o que vulgarmente se designa por um **objeto conceitual** ou **objeto semântico**".¹¹¹

¹⁰⁹ K. Thibodeau apud FERREIRA, Miguel. 2006, p.21.

¹¹⁰ K. Thibodeau apud FERREIRA, Miguel. 2006, p.21.

¹¹¹ FERREIRA, Miguel. 2006, p.23.

Os objetos conceituais assumem formas ou concepções familiares aos seres humanos, isto é, formas que existem no mundo real e que lhes são conhecidas, como livros, filmes ou fotografias. Do ponto de vista do ser humano, o objeto conceitual constitui **aquilo que deve ser preservado**¹¹².

Dessa forma, fazemos uma separação formal, que pode ser simplificada como sendo **forma e essência**. No objeto audiovisual digital, temos ambos, o suporte em que ele se encontra, o formato e como exibir isso compõem a forma. Já a imagem formada, a mensagem percebida e a interpretação possível constituem parte da essência. Um espectador pode assistir a um mesmo vídeo no cinema, no computador, na televisão, no celular que o que ele irá perceber é igual, com diferenças apenas na experiência — sonora e visual. A experiência, da qual vamos falar mais no próximo capítulo, já não é mais aquela em que a pessoa vai ao cinema contemplar o filme, mas , ao retornar para casa do trabalho, na sala de espera de uma médico, durante alguma refeição, enquanto dirigem.

Não obstante, além dos hábitos extremamente individuais, cada pessoa faz também uma interpretação particular do conteúdo que recebe. O autor chama essa compreensão de '**objeto experimentado**' e apesar da possibilidade de captura desse objeto, "nenhuma estratégia de preservação aborda seriamente esta questão" em seu livro.¹¹³

A preservação digital, nesse contexto, está sobrecarregada com a árdua tarefa de garantir que essa comunicação e essa experiência do usuário sejam garantidas no espaço e no tempo. Para isso, é inevitável considerar que as fases acima descritas estejam "acessíveis e interpretáveis".¹¹⁴

"Se a cadeia de interpretação que permite elevar um objeto digital desde o seu nível físico até ao nível conceitual for rompida, a comunicação deixa de ser possível e o objeto perder-se-á para sempre".¹¹⁵

Analisando-se de outra maneira, "um mesmo objeto conceitual pode ser representado em diversos formatos lógicos", no caso dos vídeos, estes podem ser

¹¹² FERREIRA, Miguel. 2006, p.23.

¹¹³ FERREIRA, Miguel. 2006, p.23.

¹¹⁴ FERREIRA, Miguel. 2006, p.24.

¹¹⁵ FERREIRA, Miguel. 2006, p.24-25.

.mp4, .h264, .MOV .AVI, .WMV, dentre tantos outros. Estes, ainda, podem estar inscritos em inúmeros suportes físicos, tais como DVDs, CDs, *pendrives*, HDs externos, discos rígidos e também “na nuvem”, expressão que designa quando é feito o *upload* de um arquivo para uma plataforma virtual — mas que, sempre e em qualquer medida, traz consigo um servidor físico que suporta o armazenamento desses arquivos.

Esse esclarecimento semiótico acerca dessas camadas do arquivo digital, em especial os de vídeo, nos ajuda na compreensão dos conceitos subsequentes dessa pesquisa. O que se deve preservar, então? Tudo! Mas diante da instabilidade de alguns suportes, fica difícil confiar e apostar as fichas numa solução ou outra.

Retomemos o incêndio na Cinemateca Brasileira: se as películas que foram perdidas tivessem sido digitalizadas anteriormente, perderia-se o suporte, mas a imagem compreensível permaneceria. Apesar da película original se perder, o filme se perenizaria e ainda poderia ser acessado de tantos outros dispositivos.

Apesar dos prós e contras no que se refere aos recursos para preservação, todos, até o momento, terão inúmeros benefícios e contrapartidas associadas a eles. Por mais que a fita LTO venha sendo usada com certa frequência e desconsiderando seus custos de implementação e manutenção, basta imaginar que uma fita magnética não suportaria um campo magnetizado próximo — e isso é importante! Se porventura alguma pessoa magnetizada se aproximar desse acervo, tudo se perde.

Cabe ressaltar que, quando o projetor queima um ou dois frames de um filme, isso compromete menos de um segundo de narrativa e ainda confere “charme” e ruídos, marcas ao filme — que também contam histórias. Mas, se um arquivo digital, por qualquer motivo, se corrompe, todo o filme é perdido — sem marcas ou vestígios. E aqui é importante ressaltar que o *glitch*¹¹⁶, o dano, o erro, nunca estão ao nosso alcance de explicação — eles simplesmente acontecem no mundo digital. Uns são

¹¹⁶ *Glitch* é um termo utilizado na área da computação para dizer de um pequeno problema ou falha que impede que software e/ou hardware funcionem como deveria. O termo é também utilizado na área artística porque tal dano é capaz de produzir imagens com pixels distorcidos e cores vibrantes, dessa forma muitos artistas o utilizam como técnica, forjando o erro ou mesmo simulando.

mais graves que outros, mas a maioria é irreversível e inexplicável, portanto, seria uma tentativa bastante frustrada continuar pesquisando nessa vertente.

A tecnologia de registro e de leitura é, em muitos sentidos, ainda mais vulnerável do que os suportes. A rapidez com que as tecnologias caem em desuso caracteriza o campo do audiovisual. Os formatos mudam sem parar e os suportes — conservados em boas condições — podem sobreviver à existência industrial da tecnologia de reprodução, da qual depende o acesso a eles. Todos os arquivos enfrentam o problema da manutenção de tecnologias obsoletas, descontinuadas pelas indústrias audiovisuais.¹¹⁷

Em março de 2013, a UNESCO lançou um documento preparado pela Biblioteca Nacional da Austrália, intitulado *Guidelines for the preservation of digital heritage*, que são diretrizes para a preservação do patrimônio digital. O documento possui 170 páginas e é bastante valioso como iniciativa, trazendo estudos de caso e análises consistentes sobre os formatos em alta naquela época.

Cabe ressaltar que esse tipo de documento carece de constante atualização, dada a velocidade de transição das informações. Na verdade a própria natureza de um arquivo digital garante que sempre existirão limites à pesquisa e à reflexão.

Observemos, abaixo, duas questões provenientes desses conceitos importantíssimos para o prosseguimento deste trabalho:

Digitalização: Prática de fazer a transferência de um filme originalmente analógico para algum tipo de suporte digital. ‘Você ainda não digitalizou seu acervo?’ é a pergunta feita sistematicamente aos arquivistas por políticos, administradores e supervisores dos orçamentos dos arquivos. A pergunta e seus pressupostos são muito reveladores de sua desinformação, mas o movimento para a tecnologia digital por parte das organizações da mídia tem um efeito profundo sobre a prática arquivística, as demandas de acesso e o planejamento estratégico.

Obsolescência: A digitalização coloca o dilema da obsolescência cada vez mais rápida dos formatos, com os arquivos precisando lidar com os mistérios da preservação digital, por um lado, e por outro, com a necessidade de continuar preservando e atendendo às demandas de acesso aos ‘formatos tradicionais’, mais antigos.¹¹⁸

¹¹⁷ EDMONDSON, Ray. 2017, p. 7.

¹¹⁸ EDMONDSON, Ray. 2017, p. 4.

Dessa forma, percebe-se que temos um grande desafio em mãos. Que engloba especificidades, formatos, formas de fazer e, principalmente, a realidade brasileira de recursos materiais e humanos.

2.2 Questões relacionadas à preservação

Eloy Rodrigues, em prefácio à obra de Miguel Ferreira sobre o tema, afirma que, em relação às transformações pelas quais temos passado e o crescimento do último século — demográfico, social, "evolução do acesso à educação e às tecnologias, desenvolvimento e 'enriquecimento' de múltiplas regiões e grupos sociais [e também a] revolução tecnológica dos últimos 20 anos (em particular da informática e das comunicações)", é capaz que nossa geração forneça e registre uma maior quantidade de informação "do que toda a informação registrada pelas milhares de gerações que nos precederam".¹¹⁹

Evidente que grande parte dessa informação é produzida em suportes digitais, bem como registrada e transmitida atualmente em grandes velocidades. Cabe, então, ressaltar que, assim como nos dias passados, é impossível e talvez inútil ou irrelevante preservar e salvaguardar todo tipo de informação produzida. Mas, se ainda hoje podemos consultar um material de gerações anteriores a nós (de modo particular, dos últimos cinco séculos), se faz necessário, então, assegurar que a posteridade poderá acessar um "registro igualmente significativo e relevante da produção informativa contemporânea."¹²⁰

Por isso, a preservação digital foi colocada de forma feliz por Miguel, ponderando que: "a atividade que garante que a 'comunicação entre um emissor e um receptor é possível, não só através do espaço, como também através do tempo'".¹²¹

E continua:

¹¹⁹ RODRIGUES, Eloy. 2006, p.11.

¹²⁰ RODRIGUES, Eloy. 2006, p.11-12.

¹²¹ RODRIGUES, Eloy. 2006, p 12.

Designa-se, assim, por preservação digital o conjunto de atividades ou processos responsáveis por garantir o acesso continuado a longo-prazo à informação e restante património cultural existente em formatos digitais. A preservação digital consiste na capacidade de garantir que a informação digital permanece acessível e com qualidades de autenticidade suficientes para que possa ser interpretada no futuro recorrendo a uma plataforma tecnológica diferente da utilizada no momento da sua criação.¹²²

Ele afirma que, nos últimos 20 anos, foram feitos esforços consideráveis para a construção de uma nova base de conhecimentos nessa área. Foram projetos, ideias, conceitos e estratégias que "conduziram ao reconhecimento universal do problema e à elaboração de possíveis soluções".¹²³

2.2.1 O modelo de referência OAIS

No começo dos anos 90, vários órgãos¹²⁴ se uniram com o objetivo de construir um conjunto normativo capaz de "regular o armazenamento a longo-prazo de informação digital produzida no âmbito de missões espaciais"¹²⁵ e dessa iniciativa nasceu o modelo de referência OAIS, que significa *Open Archival Information System* e é o mais completo se considerarmos os parâmetros para a preservação audiovisual digital¹²⁶.

Basicamente, o modelo OAIS, já aprovado, funciona de forma conceitual e propõe a criação de terminologias próprias e identificação dos componentes que farão parte do sistema de informação dedicado à preservação digital. "O modelo descreve ainda as interfaces internas e externas do sistema e os objetos de informação que são manipulados no seu interior".¹²⁷

No modelo há o momento do planejamento de preservação, que pensa e define políticas para cada acervo, como veremos no último capítulo desta dissertação. "Este serviço é responsável pela monitorização do ambiente externo ao repositório e

¹²² RODRIGUES, Eloy. 2006, p. 12.

¹²³ RODRIGUES, Eloy. 2006, p. 20.

¹²⁴ Consultative Committee for Space Data Systems (CCSDS) e International Organization for Standardization (ISO)

¹²⁵ FERREIRA, Miguel. 2006, p.27.

¹²⁶ Documentação sobre esse modelo encontra-se nos anexos.

¹²⁷ FERREIRA, Miguel. 2006, p.27.

por desencadear eventos de preservação sempre que necessário."¹²⁸ Dessa forma, é dever deste escolher "estratégias de preservação a utilizar no interior do repositório, monitorar as tendências comportamentais da sua comunidade de interesse ou identificar formatos que se encontram na iminência de se tornar obsoletos".¹²⁹

2.2.2 HD, LTO e possibilidades

Segundo Inês Aisengart Menezes¹³⁰, a lenta resposta dos arquivos com relação às mudanças tecnológicas está provocando grandes desfalques em nosso patrimônio audiovisual¹³¹. Em seu texto intitulado "O digital não vai esperar"¹³², a autora coloca diversas questões envolvendo a temática dos filmes digitais. Será que os primeiros filmes nacionais gravados digitalmente podem ser, ainda, encontrados e acessados? A autora aposta que muitos já se perderam.

O tema se faz frágil porque arquivos de vídeo nativos ou digitalizados sofrem com deterioração, obsolescência dos formatos — softwares e hardwares, informações que desaparecem e questões envolvendo metadados e migração de dados. Segundo a autora, as instituições já estão habituadas e consolidadas em outros formatos, sendo difícil a "incorporação de um novo espaço de materiais, ainda mais quando seu surgimento (e desaparecimento) de diferentes suportes não é acompanhado por um formato que permita preservação"¹³³.

¹²⁸ FERREIRA, Miguel. 2006, p.27.

¹²⁹ FERREIRA, Miguel. 2006, p.27.

¹³⁰ Possui mestrado no programa de Preservação e Estratégias de Difusão Audiovisual (Heritage Studies: Preservation and Presentation of the Moving Image) na Universidade de Amsterdam (2013 - 2015). Bacharel em Comunicação Social, Habilitação Cinema, pela Universidade Federal Fluminense (conclusão em 2009). Após o mestrado, trabalhou por um ano no EYE Film Institute Netherlands, uma das principais instituições de patrimônio audiovisual na Holanda, nos departamentos Digital Access e Digital Presentation, incluindo DCP e múltiplas edições e transcodificações; (...) concepção e execução de workflow de mídias digitais, metadados, preservação e migração digital. Nascida e criada no Rio de Janeiro, possui 13 anos de experiência profissional no Brasil na indústria do audiovisual e preservação audiovisual. *Texto retirado do currículo lattes da pesquisadora. <https://www.escavador.com/sobre/7264532/ines-aisengart-menezes>. Acesso em: 23 jun. 2017.

¹³¹ MENEZES, Inês. 2016, p. 42.

¹³² MENEZES, Inês. 2016, p. 42.

¹³³ MENEZES, Inês. 2016, p. 42.

Produtoras que querem guardar seu acervo, cineastas interessados em perenizar seus próprios filmes, consumidores comuns que querem guardar os vídeos de seus celulares, a Cinemateca do MAM e o próprio MIS em BH estão sempre à procura de possibilidades. O diálogo dos arquivistas é constante na ABPA (Associação Brasileira de Preservação Audiovisual) e nos encontros como o Cine OP, que há 12 anos acontecem, anualmente, na cidade de Ouro Preto em Minas Gerais, conforme já mencionado.

Cabe lembrar que esse grupo ainda é pequeno e que ainda não há parâmetros oficiais no Brasil para a preservação e salvaguarda dos acervos digitais. As buscas por ferramentas eficazes de preservação digital sempre apontam as mesmas tecnologias: HD externo, fitas magnéticas LTO, repositórios *online* e uma ou outra iniciativa não muito eficaz como as fitas ODA¹³⁴ da Sony.

Vamos compreender melhor cada uma dessas tecnologias, suas vantagens e desvantagens.

HD externo (HDDs e SSDs): são os HDs da forma como os conhecemos e costumamos fazer backup de nossos arquivos pessoais. A diferença entre os modelos HDD e SSD se dá na velocidade de acesso e na confiabilidade do tempo de duração de cada um. Mas, de modo geral, quando o termo HD é utilizado, na verdade, o sentido deste é referente ao termo HDDs.

Eles são excelentes opções para conteúdos acessados com mais frequência, além de custos baixos, mas, em contrapartida, apresentam grande taxa de erro (leitura/escrita de dados). O seu ciclo de vida também não é facilmente determinável, o que dificulta a confiabilidade, além de suscetível a corruptibilidade dos arquivos¹³⁵.

Fita magnética LTO: é o formato dominante quando o assunto é preservação digital. LTO significa *Linear Tape-Open*, em Língua Portuguesa fita aberta linear

¹³⁴ Optical Disc Archive, um sistema da Sony de armazenamento digital. Mais em: <https://www.sonypro.com.br/pro/article/broadcast-products-sony-oda>. Acesso em 13 fev. 2018.

¹³⁵ MENEZES, Inês. 2016, p.42.

(tradução livre), funciona com um padrão aberto para uma fita magnética de dados digitais que armazena até 15 terabytes de dados em sua atual geração. O número indicador ao final indica a geração, e essa revisão do formato é atualizada, aproximadamente, a cada dois anos. Estima-se que a versão 10 da fita LTO será capaz de armazenar 120 terabytes de informação.

Parece incrível à primeira vista, mas há que se lançar um olhar mais crítico sobre essa tecnologia. A principal desvantagem está no curto ou médio prazo das fitas, não porque duram pouco, mas porque exigem migração em prazos determinados. Cada geração LTO contém capacidade de leitura e escrita invertidas, sendo capaz de ler duas gerações anteriores e escrever uma geração anterior. Isso significa que, constantemente, todos os arquivos devem ser migrados de uma geração para outra antes de se tornarem inacessíveis.

Dessa forma, se estamos na geração 7, somos capazes de ler as gerações 5, 6 e 7 e de escrever novos dados apenas nas fitas de geração 6 ou 7. Isso significa que, a cada 5 anos, é necessário rever todo o acervo e realizar novos investimentos.

As vantagens são o preço, a economia de energia, a taxa de transferência alta e com poucos erros, as possibilidades de compactação e criptografia. Já as desvantagens são a gravação de dados de forma linear, o fato de o acesso constante não ser recomendado, acesso demorado e incompatibilidade do equipamento de gravação e leitura com gerações anteriores e futuras. Além disso as fitas são altamente sensíveis ao ambiente / manuseio / condições climáticas / campos magnéticos / impulsos eletromagnéticos.¹³⁶

O armazenamento de dados e arquivos em LTO como ação isolada não representa uma ação de preservação. Somente com constante migração e manutenção — pois, assim como no analógico, manter o aparelho funcionando é fundamental, aqui funciona exatamente da mesma forma¹³⁷.

¹³⁶ MENEZES, Inês. 2016, p.42.

¹³⁷ MENEZES, Inês. 2016, p.43.

Muitas iniciativas de pesquisa já se preocupam com outro aspecto, além da integridade da informação, prevendo que

haverá uma lacuna na memória social como consequência do uso das tecnologias digitais para a produção da informação contemporânea, seja ela escrita ou audiovisual. O pressuposto é que, para minimizá-la, há que se recriar uma tradição na área dos arquivos que permita uma 'continuidade consciente no tempo' por meio de pesquisas científicas para o desenvolvimento de soluções que venham assegurar a salvaguarda do legado digital contemporâneo para as gerações futuras¹³⁸.

Nas produções audiovisuais contemporâneas, "vídeos, áudio, filmes e documentação correlata têm sido produzidos quase totalmente já em meio digital"¹³⁹. Esse tipo de informação se dá em códigos e bits e bytes. Isso faz com que os suportes analógicos percam seu espaço e os "processos de criação, circulação e fruição dos conteúdos [se instaurem] de modo fluido e quase onipresente em todas as esferas da sociedade"¹⁴⁰.

Esta digitalização da sociedade criou novos espaços, tempos e movimentos que passam a exigir novas estratégias de ação para gerir e manter tais informações digitais usáveis e compreensíveis pelo máximo de tempo possível.¹⁴¹

Essa informação nascida digital trafega em redes cada vez mais velozes e efêmeras. Preservar essas estruturas, ao menos em parte, em termos de conteúdo e ambiência tecnológica, é ainda um dos maiores desafios que os profissionais que lidam com os arquivos precisam enfrentar. O contexto tecnológico, onde a informação trafega, é constantemente alterado por um mercado ávido pelo lançamento de novas gerações de equipamentos e softwares que acabam por se tornarem incompatíveis com aqueles que lhes antecederam. A informação digital, ao ser gerada, precisa obedecer a certas normas, requisitos e padrões para que sua preservação e acesso continuado sejam alcançados pelo máximo de tempo possível.¹⁴²

As possibilidades parecem aumentar quando o assunto envolve nuvens, muitas são as plataformas disponíveis atualmente e, no próximo capítulo, vamos compreender melhor se e como elas podem ajudar no processo de preservação dos arquivos audiovisuais digitais — priorizando o ponto de vista da importância do acesso.

¹³⁸ SILVA, R. D. et al. 2016, p. 2.

¹³⁹ SILVA, R. D. et al. 2016, p.2.

¹⁴⁰ SILVA, R. D. et al. 2016, p.2.

¹⁴¹ SILVA, R. D. et al. 2016, p.4.

¹⁴² SILVA, R. D. et al. 2016, p.4.



CAPÍTULO 3

YOUTUBE, NETFLIX E
EXPERIÊNCIA DO USUÁRIO

Após a compreensão do modo de funcionamento e armazenamento dos vídeos e filmes digitais, somos capazes de apresentar novas formas de trabalhar esse tipo de conteúdo — alinhadas ao crescimento rápido das tecnologias e formas de as pessoas lidarem com esse furacão de informações e mudanças repentinas.

Sabemos que hoje já é impossível pensar a própria internet sem o vídeo — que é o formato que mais cresce, segundo uma pesquisa promovida pelo Outbrain¹⁴³, principalmente vídeos de publicidade e entretenimento. Em setembro de 2017, uma pesquisa¹⁴⁴ mostrou que 87% dos usuários de internet no Brasil assistem televisão também conectados à rede web. As mídias atuais são tão interligadas que essas informações sobre comportamento do público do audiovisual se fazem relevantes para o entendimento da necessidade de se preservar tais arquivos.

O cinema do século passado se viu ameaçado pelo surgimento da TV, mais prática e bastante acessível. As pessoas iam mudando os próprios hábitos e o cinema precisou se reinventar, diversas vezes, para não perder seu público: os filmes em 3D são um exemplo disso: somente nos cinemas era possível vê-los. O diferencial dele sempre esteve na experiência: sentar na sala escura e ter sua própria imersão na narrativa da tela.

Nesse contexto é importante utilizar um termo da comunicação que se chama "convergência das mídias". Para Santaella, esse fenômeno diz respeito à "escolha entre o que se quer acessar e assistir e a descentralização da produção de conteúdos".¹⁴⁵ Nos questionários aplicados a seguir, conseguiremos perceber isso na prática, uma vez que todo mundo pode produzir e difundir seus conteúdos¹⁴⁶.

De fato, quando se busca unir "televisão, cinema e internet, há uma verdadeira experiência hipermediática"¹⁴⁷. Para analisar acesso e consumo, nesse caso,

¹⁴³ Empresa especializada em conteúdo online com escritórios em Nova York e em quatro países europeus.

¹⁴⁴ Pesquisa promovida pela Provokers.

¹⁴⁵ SANTAELLA, Lúcia. 2003 apud REIS, Isabella Rodrigues. 2017, p. 3.

¹⁴⁶ REIS, Isabella. 2017, p. 3.

¹⁴⁷ REIS, Isabella. 2017, p. 3.

devemos considerar que "as mídias tradicionais tiveram que passar por uma reestruturação em relação à produção de seus conteúdos"¹⁴⁸.

Cinema e televisão tiveram que se adaptar a essa nova realidade, e já no início dos anos 80, por exemplo, o surgimento do videocassete permitiu que os telespectadores pudessem ter acesso aos filmes saídos do cinema direto para a tela da televisão. Além disso, permitiu que outras pessoas criassem e produzissem conteúdos independentes, que podiam ser gravados na fita cassete e reproduzidos posteriormente. Mais tarde vieram CD e DVD, trazendo a questão da tecnologia digital e depois e mais especificamente na internet, surgiram diversos serviços por streaming, responsáveis por distribuir conteúdos multimídia de qualquer pessoa em diversas plataformas online, como o YouTube, Spotify ou a própria Netflix.¹⁴⁹

Filmes na televisão já rompem um pouco com essa experiência e a internet veio para mudar mesmo o paradigma da TV. Prevalentemente programada, se o espectador, porventura, perder a hora do programa, perde uma oportunidade de assisti-lo. Nesse modelo convencional, não é possível voltar para ver o que perdeu (embora algumas transmissoras de TV a cabo já permitam voltar).

Hoje, o modo de funcionamento das TVs também já tem dado sinais claros de esgotamento de seu modelo e a internet tem tomado esse espaço na cultura das pessoas, é o que afirma a reportagem abaixo (FIGURA 11), do jornal Estadão já em 2013. E é exatamente sobre isso que vamos falar neste capítulo.



FIGURA 13 – Reportagem do Jornal Estadão
 Fonte: <http://cultura.estadao.com.br/blogs/direto-da-fonte/a-televisao-brasileira-tem-dado-claros-sinais-de-esgotamento-de-seu-modelo/> Acesso em: 06 dez 2017.

¹⁴⁸ REIS, Isabella. 2017, p. 5.

¹⁴⁹ REIS, Isabella. 2017, p. 5.

As plataformas aqui mencionadas representam uma grande facilidade na distribuição de conteúdo, principalmente os independentes. Oferece, também, possibilidade de monetização e facilita a conexão dos filmes ao público-alvo, mas “essas plataformas ainda não mostraram a que vieram”¹⁵⁰.

Plataformas digitais de disponibilização de conteúdos audiovisuais como o Youtube e o Netflix [possuem estruturas econômicas diferentes], o primeiro remunera indiretamente os produtores a partir da visibilidade dos conteúdos (tempo de acesso e submissão da audiência aos comerciais embutidos antes ou durante as exibições), enquanto o segundo expande suas relações culturais e econômicas numa perspectiva transnacional, oferecendo visibilidade e distribuição de conteúdo produzido exclusivamente para esses ambientes virtuais, configurando assim um novo espaço (...) da cultura midiática¹⁵¹.

Pretende-se, neste capítulo, explicitar o modelo funcional de ambas, priorizando e problematizando a temática do acesso. Bem como desenvolver explicações acerca do armazenamento que elas propõem, que pode ser usado para a preservação dos arquivos audiovisuais de maneira geral, em especial, os digitais.

3.1 YouTube

Fundado por Chad Hurley, Steve Chen e Jawed Karim, ex-funcionários do site de comércio on-line PayPal, o site YouTube foi lançado oficialmente sem muito alarde em junho de 2005. A inovação original era de ordem tecnológica (mas não exclusiva): o YouTube era um entre os vários serviços concorrentes que tentavam eliminar as barreiras técnicas para maior compartilhamento de vídeos na internet. Esse site disponibilizava uma interface bastante simples e integrada, dentro da qual o usuário podia fazer o upload, publicar e assistir vídeos em streaming sem necessidade de altos níveis de conhecimento técnico e dentro das restrições tecnológicas dos programas de navegação padrão e da relativamente modesta largura de banda¹⁵².

O Youtube, atualmente com pouco mais de dez anos, já teve grande crescimento. E a grande aposta de seus criadores era exatamente esta: de que, em pouco tempo, grande parte do conteúdo distribuído na internet seria de vídeo, e eles tinha razão. Alguns anos depois, já é possível perceber que esse número cresce exponencialmente. Da mesma maneira, pessoas com menor conhecimento na área

¹⁵⁰ O Dilema Digital 1, 2007, p. 4.

¹⁵¹ HOFF, Rafael. 2015, p. 2.

¹⁵² BURGESS, Jean e GREEN Joshua. 2009, p.11.

audiovisual também começaram a produzir esse tipo de conteúdo, visto que todas elas conseguem fazê-lo apenas com um smartphone.

O site *www.youtube.com* possui uma aba com números impressionantes¹⁵³: mais de 1 bilhão de usuários — e isso representa quase um terço dos usuários da internet. A plataforma também alcança um número de adultos superior ao público de qualquer canal americano de TV. Metade do número total de acessos é proveniente de dispositivos celulares. Já possui versões locais em mais de 80 países e mais de 70 idiomas, o que abrange 95% dos usuários da internet. Em 2007¹⁵⁴, os dados mostram que 65 mil vídeos eram publicados por dia na plataforma. Um ano depois, o YouTube hospedava “algo em torno de 85 milhões de vídeos, um número que representa um aumento dez vezes maior em comparação ao ano anterior”¹⁵⁵. Já em 2014 um total de 70 horas de vídeo era “subido” a cada minuto¹⁵⁶. Atualmente, o canal possui 1 bilhão de horas assistidas por dia.

O YouTube quebrou barreiras e abriu espaço para que profissionais e amadores do mundo inteiro compartilhassem conteúdos audiovisuais. É também um espaço acessível e democratiza o acesso.

Embora não seja o único site de compartilhamento de vídeos da internet, a rápida ascensão do YouTube, sua ampla variedade de conteúdo e sua projeção pública no Ocidente entre os falantes de língua inglesa o tornam bastante útil para a compreensão das relações ainda em evolução entre as novas tecnologias de mídia, as indústrias criativas e as políticas da cultura popular¹⁵⁷.

Um exemplo disso é o jornalismo amador, possibilidade impensável há poucos anos. Principalmente no Brasil, onde existem fortes meios midiáticos, as pessoas conseguem ter voz utilizando a internet e principalmente o YouTube. Exemplo disso são os registros das manifestações populares no país em 2013, as imagens amadoras nesse caso são arquivos valiosos de uma época e cultura, como nos

¹⁵³ <https://www.youtube.com/intl/pt-BR/yt/about/press/> Acesso em 06 dez 2017.

¹⁵⁴ <http://observatoriodaimprensa.com.br/codigo-aberto/you-tube-a-cnn-da-era-digital-chega-ao-brasil/> Acesso em 06 dez 2017.

¹⁵⁵ BURGESS, Jean e GREEN Joshua. 2009, p.12.

¹⁵⁶ <http://www.techtudo.com.br/dicas-e-tutoriais/noticia/2014/08/vimeo-ou-youtube-qual-site-e-o-melhor-para-publicar-ideos-na-internet.html>. Acesso em 06 dez 2017.

¹⁵⁷ BURGESS, Jean e GREEN Joshua. 2009, p.8.

apresenta a pesquisadora Inês Aisengart em seu trabalho¹⁵⁸ sobre a preservação dessas imagens.

O modo de funcionamento da plataforma é bastante simples; para o espectador, basta entrar, digitar o que procura, escolher dentre as opções e assistir. Para quem deseja colocar, basta ter um *e-mail* para cadastrar como conta, preenchendo dados básicos e acessando a plataforma com todas as suas funcionalidades.

Práticas não oficiais de preservação que estimulam o uso do YouTube têm sido promovidas pela UNESCO¹⁵⁹ desde 2003, quando foi posta a questão da proteção do patrimônio imaterial. “Um método de salvaguarda empregado pela UNESCO é o armazenamento de vídeos de patrimônio imaterial no YouTube”¹⁶⁰.

Este é apenas um exemplo de uso do YouTube, visto várias ideias diferentes que surgiram a respeito de qual seria a função do site:

seria outra moda passageira, amada pela turma da tecnologia? Uma invenção inteligente que as pessoas precisariam ser convencidas a usar? Ou uma plataforma de distribuição de mídia, como a televisão? Embora a atenção dos primeiros usuários e da imprensa de massa tenha contribuído para impulsionar o serviço, a escalada do YouTube ocorreu em meio a uma névoa de incerteza e contradição sobre para que aquilo realmente servia. A missão aparente ou declarada do YouTube foi repetidamente transformada tanto pelas práticas corporativas como por sua utilização pela audiência.¹⁶¹.

Curioso ressaltar aqui que a primeira missão do YouTube era traduzida pelo “*slogan Your Digital Video Repository* (‘Seu Repositório de Vídeos Digitais’), uma declaração que, de alguma maneira, vai de encontro à exortação atual, e já consagrada, *Broadcast yourself* (algo como ‘Transmitir-se’)”¹⁶².

¹⁵⁸ MENEZES, Inês. 2014. Preserving amateur digital records of public demonstrations in 2013/2014 in Rio de Janeiro.

¹⁵⁹ Organização das Nações Unidas para Educação, Ciência e Cultura.

¹⁶⁰ PIETROBRUNO, Sheenagh. 2013, p. 2-3.

¹⁶¹ BURGESS, Jean e GREEN Joshua. 2009, p.14.

¹⁶² BURGESS, Jean e GREEN Joshua. 2009, p.14.

3.2 Netflix

A Netflix foi criada em 29 de agosto de 1997 na Califórnia (Estados Unidos), e inicialmente vendia e oferecia o aluguel de filmes para a população americana através dos correios. A empresa tinha apenas 30 empregados e 925 títulos de filmes disponíveis para locação¹⁶³. Ao longo do tempo, a empresa foi se aprimorando cada vez mais no atendimento ao cliente, principalmente através de seu site, o que gerou muitos investimentos para os fundadores. Em 2007, a Netflix deu início a um novo e revolucionário serviço que permanece até hoje: os assinantes agora poderiam assistir aos filmes e episódios de seriados originais da televisão aberta e fechada na tela do computador por meio do *streaming*, ao invés de esperar a chegada dos produtos via correspondência. Em 2009, todo o conteúdo da empresa pode ser visto pela TV através de um cabo conversor e atualmente o serviço é oferecido para mais de 40 países e possui mais de 80 mil assinaturas¹⁶⁴.

A definição formal garante que a Netflix é “uma plataforma online de *video on demand*”¹⁶⁵, com disponibilização de conteúdo audiovisual produzido em diferentes partes do mundo, com acesso via assinatura do serviço”¹⁶⁶.

Se, por um lado, podemos pensar a Netflix uma empresa puramente comercial para entretenimento, que chegou amedrontando a TV e a internet de modo geral, por outro, podemos pensar a plataforma como substituta imediata dos DVDs das locadoras. O procedimento é o mesmo: “passeia-se entre os corredores”, procurando por algo que se queira ver. Esse processo no Netflix se dá de forma ainda mais simples, passeia-se entre os títulos apenas apertando botões do controle remoto. Escolhe-se. Diferentemente da locadora, o conteúdo se adapta à tecnologia que você está usando naquele momento. À responsividade de sua tela. À velocidade de sua conexão.

Dessa forma, a internet funciona como uma ferramenta que auxilia o consumo. E, nessa pesquisa, validaremos a possibilidade de ela ajudar também a preservação audiovisual, uma vez que, como vimos no primeiro capítulo, a cinefilia e o colecionismo são grandes parceiros da preservação

¹⁶³ <http://mundodasmarcas.blogspot.com.br/2007/05/netflix-best-way-to-rent-movies.html>

¹⁶⁴ Dados de julho de 2017.

¹⁶⁵ Sigla que significa: vídeo sob demanda.

¹⁶⁶ HOFF, Rafael. 2015, p.2-3.

A Netflix nasceu de uma necessidade de transformação do mercado, foram criativos e buscaram uma alternativa digital ao esquema das locadoras — “nessa constante necessidade criativa mercadológica, faz com que empresas busquem alternativas diferenciadas em seus produtos, como é o caso aqui”¹⁶⁷.

A Netflix oferece filmes e seriados digitalizados, com possibilidade de assisti-los em computadores, televisões **[celulares]** e vídeo games. Neste processo, **a indústria** oferece a possibilidade de acesso aos filmes, sem a necessidade de contato físico com DVDs, fitas, embalagens ou papéis usados nos trâmites de empréstimo em locações. Temos aqui um exemplo de produto que consegue se manter sem o uso da loja física e da utilização de determinados recursos materiais para existir, além de estimular o download pago, prática que vem sendo discutida em decorrência dos processos da pirataria digital¹⁶⁸.

O sistema de transmissão da Netflix funciona da seguinte maneira: o usuário paga pelo serviço mensalmente e tem à sua disposição uma lista de filmes e séries disponíveis no país onde vive. Ele consegue acessá-los, de qualquer plataforma: celulares, *tablets*, televisões e computadores — de forma que a resolução se adequa ao tamanho de sua tela. Outra adequação é a velocidade da transmissão, então, o usuário pode assistir conectada a uma rede wi-fi¹⁶⁹ de alta velocidade ou mesmo ao 4G¹⁷⁰ que não será preciso aguardar o carregamento para poder assistir, o sistema da Netflix adequa não somente ao tamanho da tela, como também à velocidade da conexão.

A capacidade de carregar o artefato consigo, a portabilidade, somada à mobilidade (estar em trânsito, no ônibus, numa fila de banco) e ainda acessar os conteúdos audiovisuais on ou off-line, reconfigura aquilo que as bibliografias mais tradicionais apontam como condições de audiência e recepção dessas narrativas. Isso porque os artefatos da cultura digital permitem o download e armazenamento de temporadas inteiras de séries e seriados, alterando a forma de consumo (antes condicionada à exibição dos episódios e seus intervalos) para a possibilidade de um acesso contínuo.¹⁷¹

¹⁶⁷ HERMAN, Letícia. 2012, p. 2.

¹⁶⁸ HERMAN, Letícia. 2012, p. 3.

¹⁶⁹ Rede wi-fi ou wirelles: internet por conexões sem fio, costuma ter maiores velocidades por estarem conectadas a um modem com internet de fibra ótica.

¹⁷⁰ Rede para acessável por dispositivos móveis, costuma ter menor velocidade e ter um limite de conexão.

¹⁷¹ HOFF, Rafael. 2015, p. 11.

Segundo La Cruz, o conteúdo audiovisual acessado pela internet, bem como tudo que vemos na contemporaneidade, passa, ainda, por "um estágio provisório da técnica, não maduro, em que usos e apropriações diversos estão sendo experimentados e, ao mesmo tempo, usos anteriores e suas interfaces são copiados e simultaneamente tensionados"¹⁷². Para ele, "a criação e o uso de hardwares e softwares de conexão, e acesso à internet, vão reorganizando a vida social em um processo que está em andamento"¹⁷³.

Outra questão importante nesse contexto é pensar as interrupções que a narrativa audiovisual sofrerá com esses avanços:

Mesmo que a área de concentração do espectador esteja voltada para uma tela de três polegadas do seu smartphone, ele está também percebendo o que acontece ao seu redor pela visão periférica, pelo tato, pelo olfato. Diferente da condição doméstica, em que o espectador tem maior controle sobre as variáveis (pode fechar a porta, aumentar o volume do som, fechar a cortina da janela), num ônibus o processo de imersão na narrativa é condicionado pelo convívio social (algumas leis coíbem a audição de conteúdos sonoros por autofalantes de dispositivos móveis no transporte coletivo urbano brasileiro e exigem o uso de fones de ouvido, por exemplo). Quanto dessas interferências interrompem, influenciam ou mesmo reconfiguram essa recepção ou consumo? Apesar de reconhecer a diferença entre essa recepção e aquela ambientada no espaço doméstico ou do cinema, poucos estudos científicos têm investigado tais correlações¹⁷⁴.

Você pode estar se perguntando o que isso tem a ver com preservação, afinal. Trata-se de um novo padrão de distribuição, que traz consigo uma grande acessibilidade. A internet trabalha com possibilidades. "Por conta disso, entendemos que o processo de digitalização surge como uma nova alternativa para enfrentar as marcas do tempo, a seleção do mercado, os efeitos do esquecimento"¹⁷⁵. Mesmo assim é impensável conceber esta como uma solução definitiva. "A informação digitalizada também tem suas vulnerabilidades, pois depende, por exemplo, da eficácia dos sistemas de informática e dos softwares"¹⁷⁶, e esta é apenas uma dentre tantas questões e dilemas que envolvem o digital.

¹⁷² LA CRUZ, Sonia. 2012, p. 14.

¹⁷³ LA CRUZ, Sonia. 2012, p. 14.

¹⁷⁴ HOFF, Rafael. 2015, p. 11.

¹⁷⁵ MUSSA, Ivan; et al. 2012, p.5.

¹⁷⁶ MUSSA, Ivan; et al. 2012, p.5.

Um serviço como a Netflix, nos EUA, por exemplo, mantém seu foco nas produções audiovisuais de grande orçamento, mas por serem alcançados digitalmente, em tempo real, a atenção também se desloca para filmes antigos e/ou de menor expressão. Um exemplo ainda mais claro disso é o site Mubi.com. Ele opera de maneira parecida com a do Netflix, disponibilizando-os para serem assistidos online. O Mubi, porém, privilegia obras independentes e/ou de fora do eixo hollywoodiano. Películas dos mais variados países (Argentina, Rússia, Coreia do Sul, etc.) fazem parte do acervo. Na rede de exibição cinematográfica, produtos como esses encontrariam mais dificuldades para achar seu público, apesar de existir grande demanda. O mercado digital proporciona a chance de selecionar e agregar os possíveis consumidores espalhados pelo mundo, graças a serviços como estes (...) a internet favorece esse tipo de estratégia de negócio.¹⁷⁷

O Mubi.com é só mais um exemplo, de plataforma semelhante à Netflix e traz consigo a grande vantagem do acesso a filmes que não teriam essa possibilidade exatamente por estarem fora do eixo de consumo e distribuição hollywoodana.

O YouTube e todos os portais de vídeo on-line que o seguiram transformaram definitivamente a nossa maneira de absorver conteúdo. O momento agora não é de aguardar o próximo YouTube, [ou qualquer outra plataforma], mas de descobrir o que essas ferramentas farão daqui para a frente e o que poderemos fazer com elas.¹⁷⁸

3.3 Experiência dos usuários – acesso aos filmes

A preservação não é um fim em si. Ela é necessária para assegurar o acesso permanente e careceria de sentido sem esse objetivo. (...) Além disso, a natureza relativamente frágil e efêmera dos suportes e das tecnologias audiovisuais, assim como as restrições jurídicas e comerciais que pesam sobre o acesso, colocam essas funções no centro da gestão e da cultura dos arquivos audiovisuais. Podemos dizer que a preservação abarca o conjunto de operações necessárias para perenizar o acesso a documentos audiovisuais no maior grau de sua integridade. Ela pode englobar um grande número de procedimentos, de princípios, de atitudes, de equipamentos e de atividades. A preservação engloba, por exemplo, a conservação e a restauração de suportes; a reconstituição de uma versão original; a copiagem e o processamento do conteúdo visual e/ou sonoro; a manutenção dos suportes em condições adequadas de armazenamento, a recriação ou imitação de procedimentos técnicos obsoletos, de equipamentos e de condições de apresentação; a pesquisa e a coleta de informações para conduzir a bom termo todas essas atividades. A palavra acesso possui igualmente um amplo leque de significados. Ela designa qualquer forma de utilização das coleções, dos serviços ou dos conhecimentos de um arquivo, (...). O acesso pode ter um caráter ativo (de iniciativa da própria instituição) ou passivo (de iniciativa dos usuários da instituição).¹⁷⁹

¹⁷⁷ MUSSA, Ivan; et al. 2012, p.7.

¹⁷⁸ BURGESS, Jean e GREEN Joshua. 2009, p.6.

¹⁷⁹ EDMONDSON, Ray. 2013, p. 76-78.

Ray Edmondson é sincero ao propor o que se defende nesta dissertação, de que a preservação não termina em si mesma e que o acesso é a outra face dessa moeda. Segundo ele, por comodidade, tratam-se esses conceitos de forma separada, mas nunca devemos esquecer de que eles "guardam entre si uma relação de interdependência, pois o acesso é parte integrante da preservação." ¹⁸⁰ Prover acesso é, portanto, a razão de ser do arquivo.¹⁸¹

O digital democratiza o acesso: esta é outra premissa básica desta dissertação. Sabe-se que pela internet os arquivos podem chegar a qualquer canto do mundo em pouquíssimo tempo. Os modos de consumo do mundo contemporâneo já demonstram que as pessoas se habituaram a consumir o audiovisual por meios e dispositivos digitais, e, como veremos adiante, elas já nem possuem o hábito de comprar um DVD de filme que gostam, mas estão dispostas a pagar mensalmente por um acesso ilimitado a conteúdo audiovisual.

A internet como rede mundial é, nos dias atuais, uma importante ferramenta de serviços e pesquisa, porém, "coloca também novas questões relativas à seleção, incorporação e preservação, bem como questões de ordem conceitual e ética" ¹⁸² que não serão aprofundadas nesta dissertação. De qualquer forma, ressalta-se que "os arquivos devem, mais ainda do que as outras instituições de preservação, articular sua política de acesso levando em conta as realidades comerciais ligadas ao direito de autor". ¹⁸³ E os desafios éticos são muitos, desde as possibilidades de política para acesso até o pensamento sobre a manipulação digital, que pode subverter a própria história. São, portanto, muitos os dilemas e pressões éticas.¹⁸⁴

Para além do campo cinematográfico, as práticas sociais hoje são medidas por dispositivos audiovisuais. Informações criadas e disseminadas em redes sociais, por exemplo, são manifestações culturais e estão fora do âmbito de coleta das

¹⁸⁰ EDMONDSON, Ray. 2013, p. 76.

¹⁸¹ EDMONDSON, Ray. 2013, p. 39.

¹⁸² EDMONDSON, Ray. 2013, p. 40.

¹⁸³ EDMONDSON, Ray. 2013, p. 79.

¹⁸⁴ EDMONDSON, Ray. 2013, p. 40.

instituições arquivísticas tradicionais, dessa forma, corre-se o risco de esvaecimento.¹⁸⁵

Uma pesquisa **qualiquantitativa** foi feita para que pudéssemos ter uma noção de como as pessoas lidam com seus próprios arquivos. O nicho selecionado foi o de pessoas que trabalham com tecnologia, por estarem habituadas aos termos, por pressupormos que elas já possuem uma ligação com backups e nuvens, bem como consomem conteúdo audiovisual via *streaming*.

3.3.1. Metodologia e resultados

A ferramenta utilizada para a formatação das perguntas do questionário foi o Google Forms¹⁸⁶, que permite a criação de questões abertas e fechadas e fornece os resultados em forma de gráficos e tabelas. A escolha se deu em virtude da facilidade de coletar as respostas, num sistema em que é possível mandar um *link* e o público consegue responder *online* e de qualquer dispositivo.

No apêndice estão as perguntas aplicadas cujos resultados serão aqui discutidos. Essas informações serão muito úteis ao capítulo final, que propõe algumas ações possíveis aos arquivos.

Esperava-se que essas pessoas já compreendessem, minimamente, que a preservação audiovisual se faz relevante. O resultado, no entanto, não foi este. Revelou-se que mesmo estas pessoas — mesmo acostumadas com tecnologia — ainda perdem seus arquivos e não compreendem a importância de se ter arquivos preservados para a posteridade.

Não foi pedida a identificação dessas pessoas no questionário por questões éticas, e porque os hábitos delas são mais relevantes para o propósito deste trabalho.

¹⁸⁵ MENEZES, Inês. 2016, p. 42.

¹⁸⁶ docs.google.com/forms/

A pesquisa foi aplicada em uma startup¹⁸⁷ de produtos digitais¹⁸⁸ de Belo Horizonte. A escolha por esse nicho priorizou o fato de que essas pessoas estão habituadas aos termos utilizados e possuem um conhecimento mesmo que mínimo sobre a própria web e formatos dos arquivos de áudio e vídeo.

A escolha foi feita também por eles estarem no meio exato entre "leigos" que nunca ouviram falar desse assunto e não conhecem a web ou estrutura dos arquivos tão bem (a maioria da população de nosso país) e a outra ponta que são os especialistas, arquivistas, cineastas, funcionários de bibliotecas e arquivos, conservadores e restauradores ou entusiastas que pesquisam sobre o tema, eles são minoria absoluta e para eles não ficaria dúvida alguma da necessidade de se preservar.

Num montante de aproximadamente 200 pessoas, praticamente 80% delas responderam ao questionário de maneira anônima e sincera. A análise dos dados coletados a seguir será de grande valia para entendermos como esse público lida com seus próprios arquivos na era digital e o que esperar do público leigo em relação ao assunto, bem como dos especialistas e funcionários de arquivos.

Para a compreensão dos resultados abaixo, deve ser feita uma margem de erro de aproximadamente 5% para mais ou para menos. Isso porque se a pessoa, porventura, marca por engano alguma opção errada, ela não consegue voltar e refazer o questionário, pois o sistema somente coleta as respostas uma vez por pessoa.

¹⁸⁷ Empresa, geralmente pequena, de base tecnológica, cujo modelo de negócio é diferenciado e trabalha geralmente com ideias criativas e inovadoras, usando a internet como impulsionadora.

¹⁸⁸ Cursos online e conteúdos em diversos formatos: áudio, vídeo, e-books, slides e afins.

Qual a sua faixa de idade?

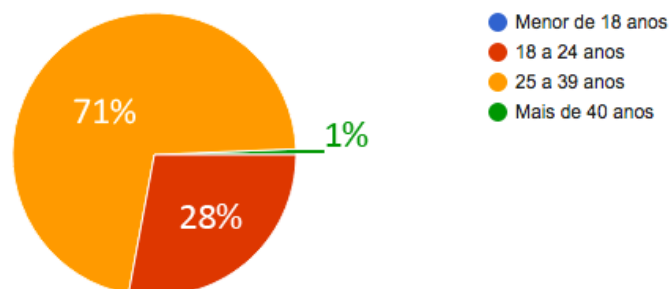


FIGURA 14 – Gráfico do questionário de pesquisa: faixa etária. Gerado pelo Google Forms.

A faixa etária dos respondentes está predominantemente entre 25 e 39 anos (71%) e faz parte de uma geração que conviveu com o analógico na infância e foi se acostumando ao digital. Em seguida, estão os jovens e entre 18 e 24 anos (28%), que já nasceram conhecendo mídias digitais e pouco conviveram com o analógico. Os outros estão acima dos 40 (1%), uma geração que costuma ter maior dificuldade com o digital.

O nível de escolaridade também varia: quase 50% dos respondentes possuem nível superior completo. 21% possuem pós-graduação completa, 11% estão concluindo uma pós graduação. 13% ainda estão fazendo a primeira faculdade e 6,4% largaram os estudos acadêmicos. 2% dos entrevistados ainda estão no ensino médio e/ou técnico profissionalizante. 2% dos entrevistados ainda estão no ensino médio e/ou técnico profissionalizante.

Qual o seu nível de escolaridade?

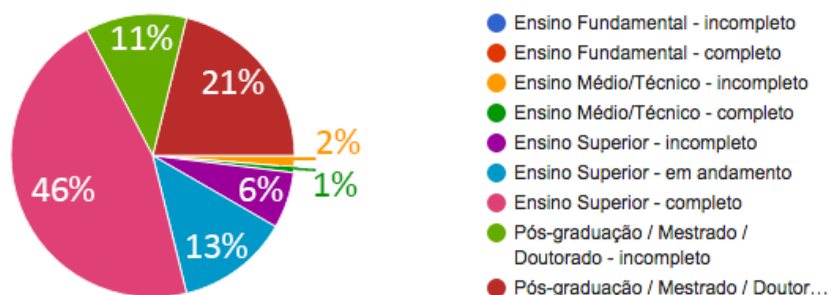


FIGURA 15 – Gráfico do questionário de pesquisa: escolaridade. Gerado pelo Google Forms.

Cabe ressaltar que os resultados demonstram uma grande preocupação com estudo. Visto que, somados, 70% possuem curso superior e também somados, a busca por pós-graduação incluindo quem ainda está fazendo está em 31%.

Em relação a todos os entrevistados, a maioria absoluta trabalha, visto a escolha do local onde a pesquisa foi aplicada. Os itens estágio, meio horário, *freelancer* e autônomo designam categorias diferentes de trabalho. Os 2% que alegam não trabalhar foram computados como resposta inválida ou preenchimento incorreto, uma vez que todos os respondentes trabalham.

Você trabalha?

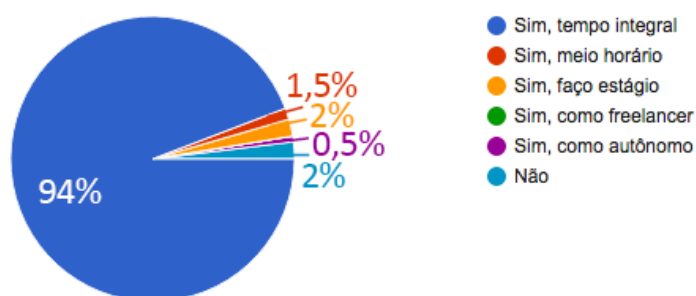


FIGURA 16 – Gráfico do questionário de pesquisa: emprego. Gerado pelo Google Forms.

As profissões que mais aparecem são: desenvolvedores (programadores *java*, *front-end* e *back-end*), analistas de suporte e marketing, redator, *designer* e *videomaker*. Nessa primeira parte, podemos ter apenas uma ideia do público com o qual estamos lidando: pessoas que estudaram e estão se qualificando cada vez mais em suas áreas específicas.

Questões socioeconômicas simples apresentadas nesta dissertação e não aprofundadas são importantes para a compreensão desses resultados, visto que o acesso à internet, a possibilidade de ter TV a cabo em casa, assinar Netflix, possuir dispositivos diversos como celular, *tablet*, notebook, computador, televisão constituem fatores relevantes que influenciam nos resultados.

A segunda parte do questionário trata dos hábitos de consumo, para entender como essas pessoas se comportam. 80% dos entrevistados afirmam consumir muito conteúdo audiovisual, 18% afirmam consumir pouco e 2% afirmam não consumir.

Você tem o hábito de consumir conteúdo audiovisual? (filmes, séries e afins...)

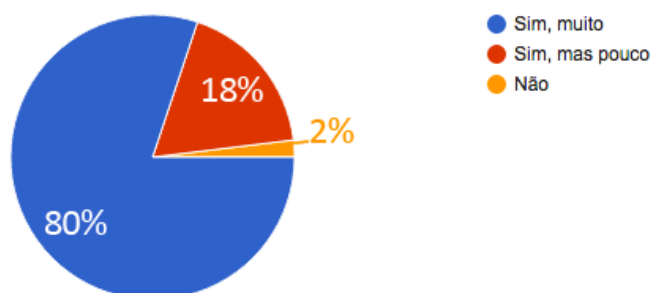


FIGURA 17 – Gráfico do questionário de pesquisa: consumo audiovisual. Gerado pelo Google Forms.

Esse já era um resultado esperado, pois sabe-se que a grande maioria das pessoas gosta de assistir a produtos audiovisuais, como filmes e séries. Dentre os maiores hábitos, estão: assistir filmes e séries na TV 70%, assistir a filmes e séries em outros dispositivos (computador, *tablets* e celulares) 68%, ir ao cinema 67% e participar de mostras e festivais apenas 9%.

Marque as alternativas que melhor representam seus hábitos (se houverem):

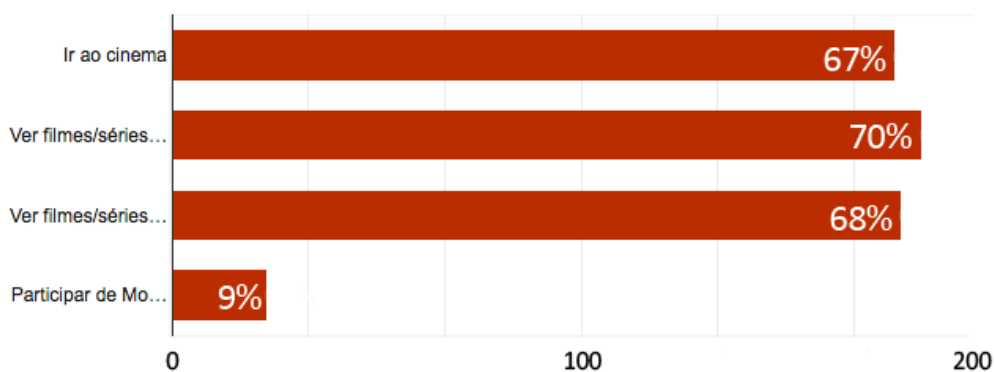


FIGURA 18 – Gráfico do questionário de pesquisa: hábitos de consumo. Gerado pelo Google Forms.

Outra questão relevante a essa pesquisa é o hábito de baixar filmes na internet, que, antes da expansão das duas plataformas acima mencionadas, era bastante comum. Importante ressaltar que esse hábito é, na maioria das vezes, ilegal. O avanço de plataformas como o próprio Netflix regulariza essa questão e, embora não tenha todos os filmes sempre à disposição e tenha uma busca otimizada por país, já ajuda a evitar e reduzir esse hábito, que, como podemos ver abaixo, ainda é bastante comum. Estes, “que sempre fazem” somados aos que fazem apenas “de vez em quando” significam um total de 60% das pessoas ainda baixando filmes por métodos não autorizados.

Você baixa filmes na internet? (torrent, etc.)

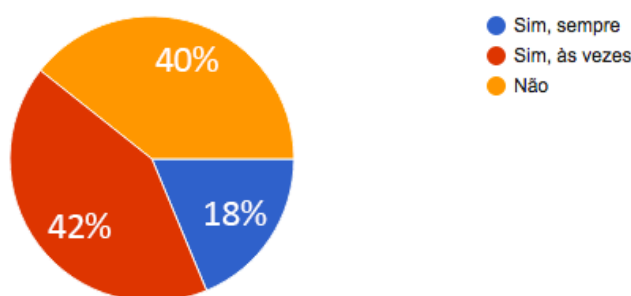


FIGURA 19 – Gráfico do questionário de pesquisa: download de filmes. Gerado pelo Google Forms.

Outro ponto de destaque é o pagamento de TV a cabo. Mais de 60% dos entrevistados ainda possuem o serviço, o que significa que a internet chegou, mas ainda não fez mudar os hábitos do consumidor a ponto de ele abrir mão de sua programação de TV.

Você assina TV paga? (sky, net, claro, etc.)

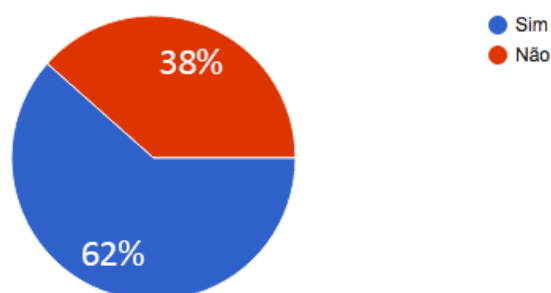


FIGURA 20 – Gráfico do questionário de pesquisa: assinatura de TV paga. Gerado pelo Google Forms.

Da mesma forma, temos também uma maioria absoluta que já aderiu ao Netflix. Uma razão para isso refere-se à qualidade e à variedade do serviço prestado, além da taxa bastante competitiva e valor muito abaixo das TVs pagas.

Você assina Netflix?

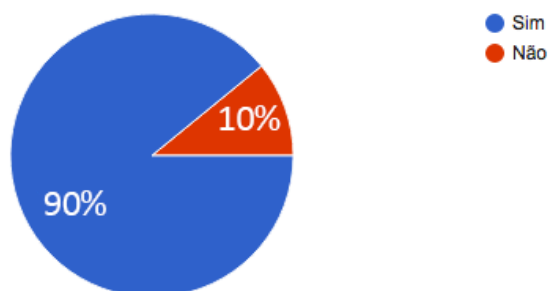


FIGURA 21 – Gráfico do questionário de pesquisa: assinatura Netflix. Gerado pelo Google Forms.

E, como estamos falando de internet, quase metade dos entrevistados já utilizou o YouTube para assistir filmes de longa metragem. O número impressiona e, pela forma como temos visto, tende a crescer.

Você já assistiu filmes (longa duração) pelo YouTube?

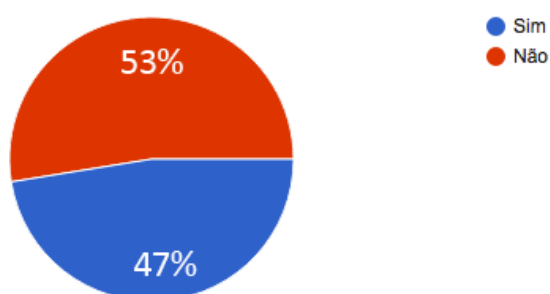


FIGURA 22 – Gráfico do questionário de pesquisa: filmes no YouTube. Gerado pelo Google Forms.

Também foi perguntado se essas pessoas já pagaram por algum serviço pontual de *streaming* por demanda. Funciona também como uma locadora na qual você paga um valor por filme, e não como uma mensalidade que te dá direito a assistir a todas as coleções disponíveis. Percebe-se que essa prática já não é tão comum.

Já contratou filmes por demanda? (Net NOW, Google Play, Vivo Play, Sky Online, Mubi, etc...)

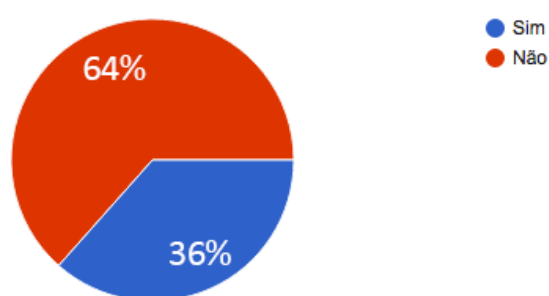


FIGURA 23 – Gráfico do questionário de pesquisa: filmes por demanda. Gerado pelo Google Forms.

O questionamento a seguir diz respeito também ao que esse trabalho quer defender. A internet é um espaço infinito, que pode ser amigo da preservação. Nela encontramos todo tipo de conteúdo que nos interessa, sabendo onde e como procurar. Quando uma pessoa diz que não encontrou algum filme, isso deve ser levado em consideração. Alguns filmes raros são encontrados apenas em cinematecas e arquivos, mas seria incrível se, porventura, fossem digitalizados e acessados facilmente online. Até porque, sequer teríamos a possibilidade de estudar cinema atualmente se os filmes dos primórdios não estivessem hoje acessíveis, como já vimos no primeiro capítulo.

Já ficou sem encontrar algum título que desejava na internet?

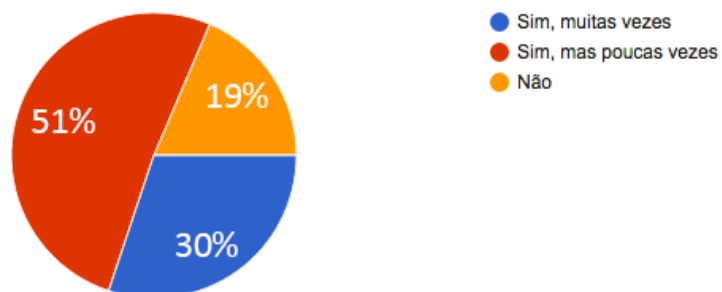


FIGURA 24 – Gráfico do questionário de pesquisa: títulos não encontrados. Gerado pelo Google Forms.

A questão abaixo também toca diretamente nos assuntos tratados no primeiro capítulo, a respeito do colecionismo. As pessoas já não estão pagando mais por filmes e programas favoritos. Isso pode estar relacionado diretamente ao fato de que elas já pagam a TV a cabo e o serviço de *streaming* para assistirem ao que gostam.

Você compra DVD ou Blu-ray dos filmes/séries que gosta?

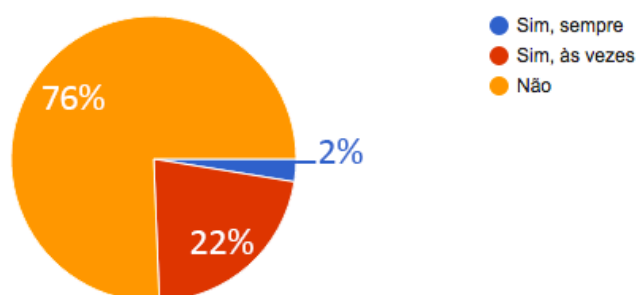


FIGURA 25 – Gráfico do questionário de pesquisa: compra de DVDs e Blurays. Gerado pelo Google Forms.

Abaixo, outra questão importante. Já sabemos quem é nosso público, mas importante ressaltar que 21% dos entrevistados não consideram importante preservar e guardar, e 30% sequer têm opinião formada sobre o assunto. Talvez eles não saibam que, para que possam assistir aos filmes e a séries que tanto gostam, seja preciso fazer um trabalho de preservação. Este é um ponto crítico para percebermos que, talvez, em relação ao público leigo, esses números fossem ainda maiores.

Você considera importante preservar e guardar filmes?

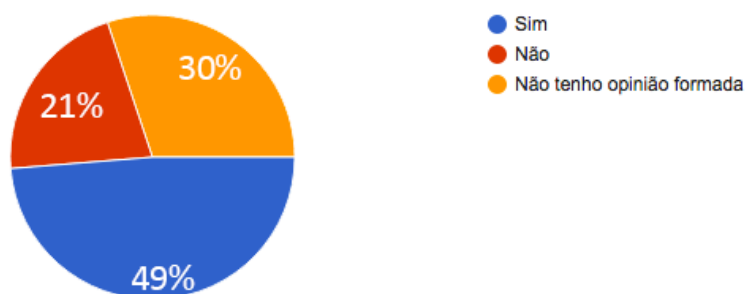


FIGURA 26 – Gráfico do questionário de pesquisa: importância preservação. Gerado pelo Google Forms.

No terceiro momento da pesquisa a ideia era colocar o espectador também no lugar do produtor de conteúdo, para que, focado em suas próprias experiências, seja capaz de compreender a questão. Hoje em dia sabemos que todo mundo é capaz de tirar fotos e gravar vídeos com o próprio celular. Apenas 30% dos entrevistados afirmou não produzir nenhum tipo de conteúdo, o que representa um número alto, visto que o que se observa é um aumento exponencial de produção de fotos e vídeos, principalmente para redes sociais.

Você produz conteúdo audiovisual? (Fotos, vídeos, áudios).

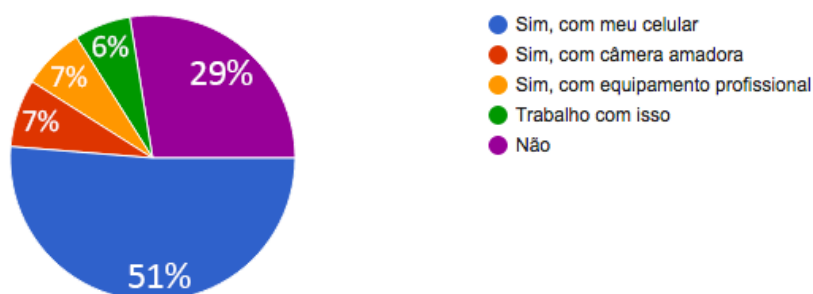


FIGURA 27 – Gráfico do questionário de pesquisa: produção conteúdo audiovisual. Gerado pelo Google Forms.

Ao serem questionados sobre o compartilhamento desses conteúdos audiovisuais, mais de 60% dos entrevistados afirmou compartilhar parte deles e 7% todos eles.

Você tem o hábito de postar/compartilhar esses conteúdos?

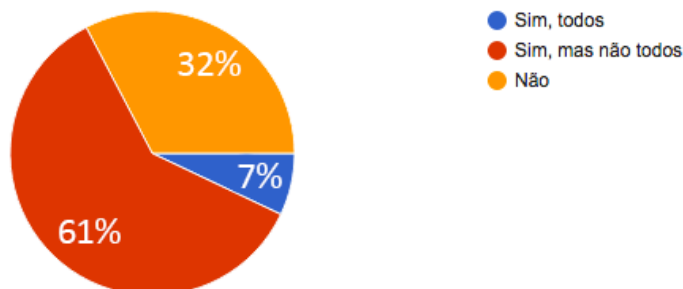


FIGURA 28 – Gráfico do questionário de pesquisa: hábitos de postagem. Gerado pelo Google Forms.

E as redes sociais que escolheram para isso são muitas. Em primeiro lugar, está o Instagram com 68%, em segundo lugar, o Facebook, com 58%. Seguidos pelo Whatsapp, Twitter, Youtube, Vimeo, Google Plus, Snapchat, Pinterest, Behance, Flickr, 500px e Sparkle, dentre outros.

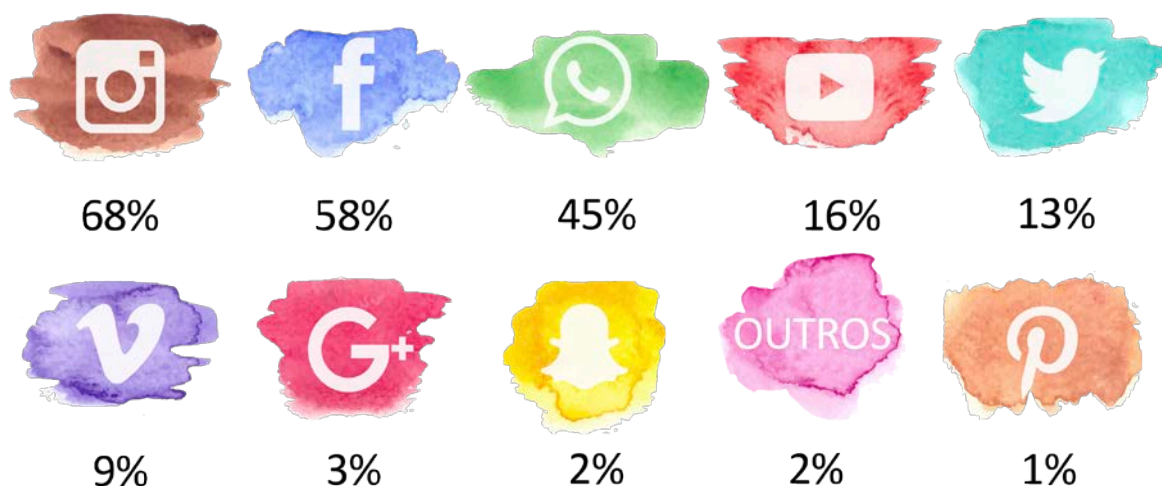


FIGURA 29 – Gráfico do questionário de pesquisa: porcentagem redes sociais. Produzido pela autora com os dados do Google Forms.

Compartilhar é importante, mas não garante perpetuidade desses arquivos. Para tanto, foram questionados a respeito dos seus hábitos de backup. Apenas 13% afirmaram não fazerem/nunca terem feito backup.

Você tem o hábito de fazer backup em nuvem? (Google Drive, Dropbox, Mega, iCloud, One Drive, etc.)

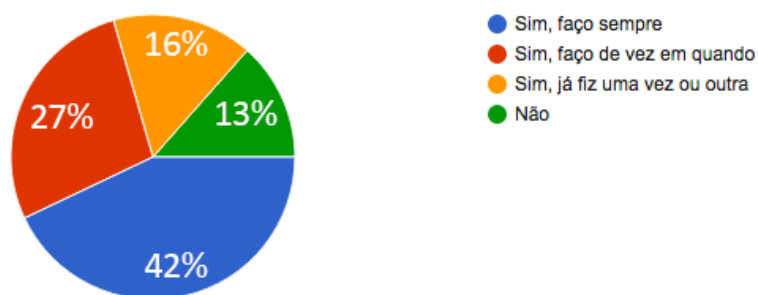


FIGURA 30 – Gráfico do questionário de pesquisa: backup em nuvem. Gerado pelo Google Forms.

Para compreender ainda melhor os que têm o costume de fazer backup, foram questionados também onde os arquivos pessoais eram guardados. Em primeiro lugar, está o próprio computador, seguido do backup em nuvem. Os menores índices foram para pendrives e HDs externos, cartões de memória, armazenamento em plataformas de compartilhamento e backup pelo próprio *e-mail*.

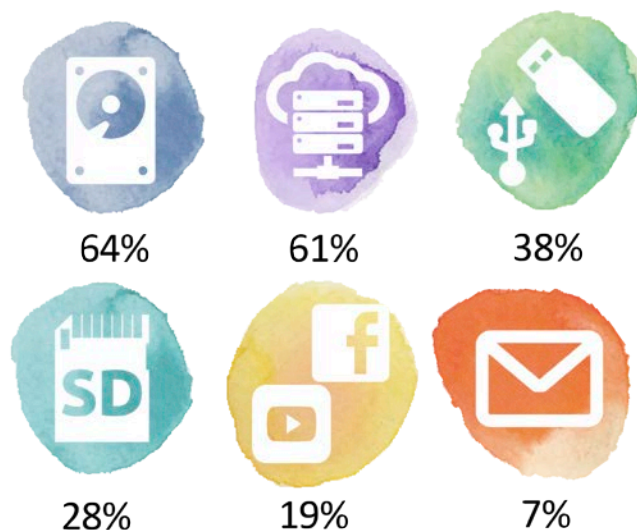


FIGURA 31 – Gráfico do questionário de pesquisa: backup em dispositivos físicos. Produzido pela autora com os dados do Google Forms.

Ao serem questionados sobre a segurança dos arquivos na internet, as pessoas demonstraram confiar bastante. E, mesmo receosos ou com ressalvas, colocam seus arquivos pessoais na rede.

Você confia na segurança dos arquivos na internet?

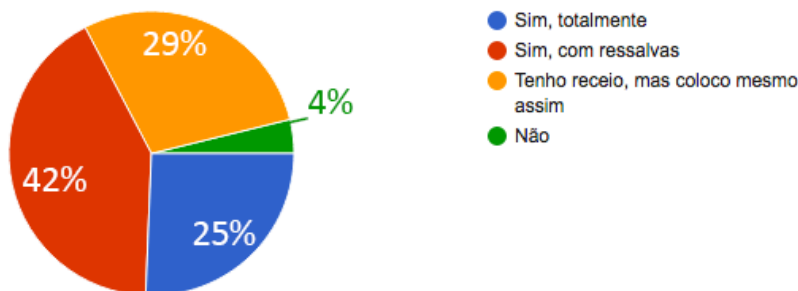


FIGURA 32 – Gráfico do questionário de pesquisa: confiabilidade na web. Gerado pelo Google Forms.z

Por fim, as pessoas foram convidadas a observarem as imagens abaixo e responderem se percebiam nelas alguma diferença visual de resolução.



IMAGEM 1



IMAGEM 2

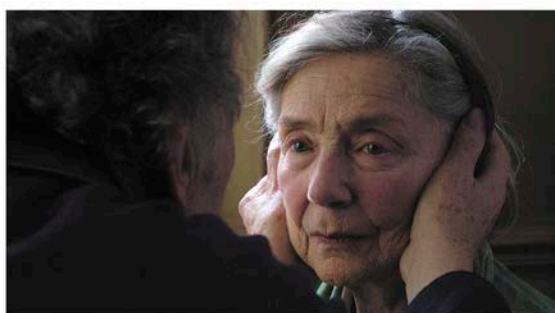


IMAGEM 1

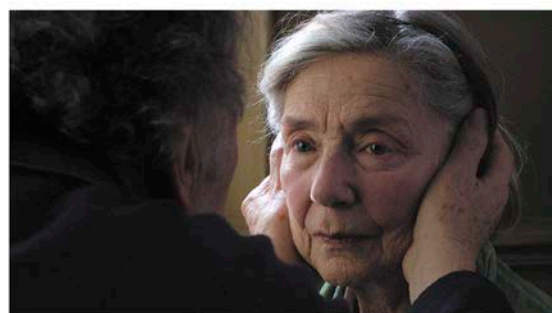


IMAGEM 2

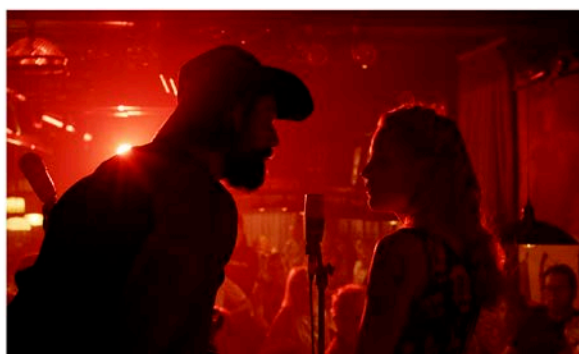


IMAGEM 1

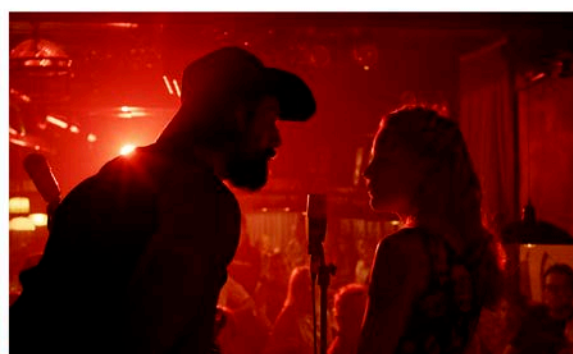


IMAGEM 2

FIGURA 33 – Gráfico do questionário de pesquisa: comparativo de imagens. Produzido pela autora.

Em todas as opções, há uma imagem em altíssima resolução (4k – 4096x2160 pixels) e outra que passou por uma compressão que a transformou em Full HD (1920x1080) por meio do Photoshop.

As respostas para as percepções dessas questões foram muito divergentes, a grande maioria das pessoas considerou que ambas eram iguais em qualidade. A sequência feita foi: qualidade superior nas imagens 1 da primeira comparação, 2 da segunda comparação e 1 da terceira. As respostas estão a seguir:

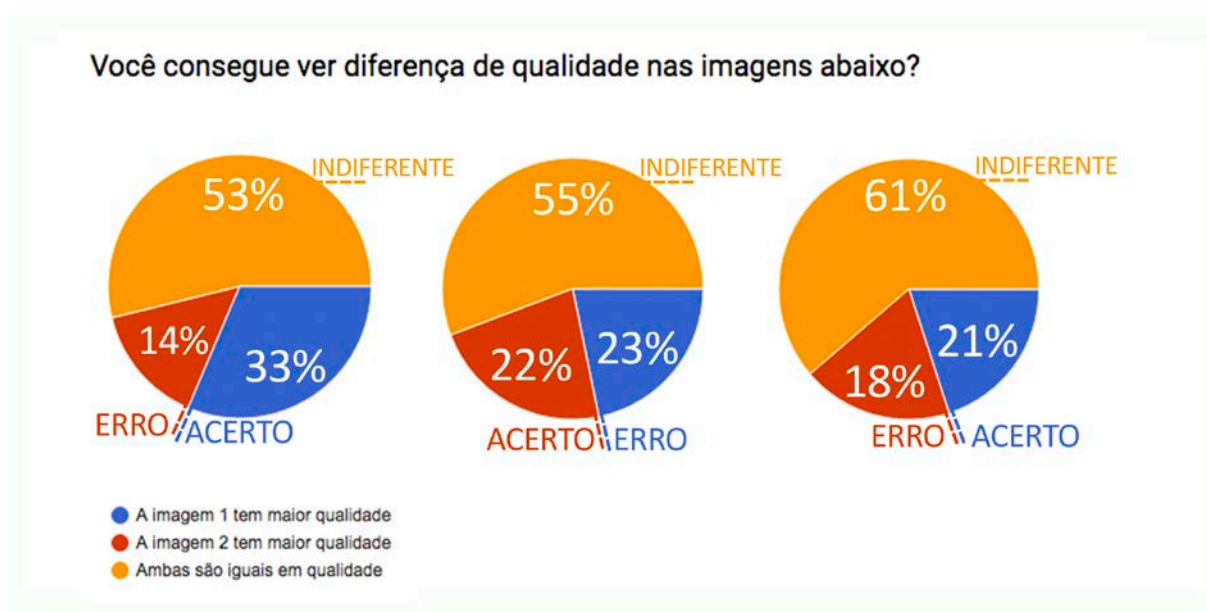


FIGURA 34 – Gráfico do questionário de pesquisa: percepção de qualidade. Gerado pelo Google Forms.

Para a grande maioria, nos 3 casos, a diferença não foi perceptível. E os acertos variaram entre 30% e 20%. Além disso, pessoas que viram qualidade onde de fato não havia, com taxas de erro entre 14% e 20%.

Essa questão revela o fato, que, talvez, muitos não tenham percebido, de que as imagens vistas em telas de celulares, *tablets*, *notebooks* e mesmo das televisões são imagens que passam por processos de conversão com perda pouco significativa na sensação de qualidade.

Um dos entrevistados, posteriormente à aplicação dos questionários procurou-nos para dar um depoimento: “tenho todas as minhas fotos do celular salvas no Google Fotos desde 2013, quando tive meu primeiro celular Android”. E, para isso, ele não precisou sequer se esforçar, ativou a ferramenta que atualiza automaticamente e já faz o backup das imagens do celular na nuvem. O dispositivo móvel é passível de estragos, pode queimar, ser roubado, ou mesmo perdido, que as fotos estarão lá.

Porém, tal declaração não é suficiente para salvaguarda de arquivos. Afinal, ferramentas como essas acabam se tornando obsoletas e/ou são descontinuadas com frequência pelo próprio fabricante. De fato, teremos problemas se confiarmos a preservação dos nossos arquivos pessoais a terceiros, mas ainda não existem opções que descartem essa possibilidade por completo. Mesmo os grandes cinemas lidam com dificuldades em seu próprio acervo audiovisual digital. Produtores independentes possuem a mesma questão ainda mais acentuada pela falta de recursos. A seguir, apresentaremos algumas iniciativas de baixo orçamento para pessoas e instituições.



CAPÍTULO 4

E O QUE PODE SER FEITO?

As iniciativas e pesquisas no âmbito da preservação audiovisual digital ainda não foram suficientes para declarar uma maneira eficaz de preservação dos arquivos de vídeo e filmes digitais, sejam eles profissionais ou amadores. Fato é que não sabemos bem sequer o que iremos fazer para armazenar e preservar nossos próprios arquivos.

Seria muita ousadia trabalhar com o termo “soluções”, mas antes, traremos os termos “propostas” e “amostras”. Dessa forma, este capítulo servirá como um “dossiê” sucinto, que cataloga algumas das maneiras testadas até então, trazendo informações do que tem sido feito com esse propósito.

Os principais métodos recomendados para a preservação dos objetos digitais podem ser agrupados em dois tipos: os estruturais e os operacionais¹⁸⁹:

os estruturais são os investimentos iniciais das instituições que estão se preparando para implementar algum processo de preservação e que adotam ou adaptam um dos modelos de metadados existentes. Já as atividades operacionais são as medidas concretas aplicadas aos objetos digitais.¹⁹⁰

Neste último capítulo, iremos mencionar elaborações que contemplam essas duas vertentes e podem ajudar pessoas e instituições na elaboração de uma metodologia de preservação audiovisual digital que atenda às expectativas e necessidades.

Uma grande e pioneira iniciativa registrada nesse sentido se deu antes mesmo da série de livros “Dilema Digital”. Em 2004, foi lançado pela Academia Americana de Artes e Ciências Cinematográficas o livro "*Long-term management and storage of motion picture materials*"¹⁹¹ que trazia uma introdução sobre o assunto e estudos de caso de Hollywood.

Já em 2007 foi lançado o "Dilema Digital — questões estratégicas na guarda e no acesso a materiais cinematográficos digitais", também pela Academia. Este aponta “as crescentes preocupações com a longevidade dos materiais cinematográficos

¹⁸⁹ ARELLANO, 2004 apud Política de Preservação – PINACOTECA. 2017, p.4.

¹⁹⁰ Política de Preservação – PINACOTECA. 2017, p.4

¹⁹¹ Gerenciamento e armazenamento a longo prazo de materiais cinematográficos (tradução livre).

digitais” e explora tais questões considerando não apenas o meio audiovisual, mas também as áreas tantas de pesquisa que padeciam do mesmo problema: “a falta de garantia de acesso de seus dados digitais a longo prazo”¹⁹².

O livro traz a questão deste trabalho, ao declarar que "as atividades que envolvem dados digitais geram um volume muito maior de materiais e, atualmente, uma parcela muito pequena desse volume é preservada"¹⁹³. O autor, ainda, nos lembra de que "a estratégia do 'salvar tudo' foi possível devido ao baixo custo de armazenamento e à longa duração da película fotoquímica e da tecnologia a ela associada"¹⁹⁴, o que levanta novas questões acerca do custo de preservação de acervos digitais:

a explosão da informação digital vem acompanhada da ameaça de extinção dessa mesma informação e, por conseguinte, da perda de conteúdos valiosos. Se um único filme digital gera mais de dois petabytes de informações digitais — o equivalente a quase meio milhão de DVDs —, as decisões sobre quais materiais guardar, quais preservar e quais ações de gerenciamento de risco são necessárias antes da determinação da migração trazem novas pressões para esse gerenciamento.¹⁹⁵

O Dilema Digital 2 foi lançada anos mais tarde, em 2012, com o objetivo de aprofundar ainda mais as questões do primeiro e trouxe um enfoque diferente: “o impacto das tecnologias digitais nas comunidades de cineastas independentes e de Arquivos audiovisuais sem fins lucrativos”¹⁹⁶ e a perspectiva dessas duas realidades — que se aproximam da brasileira, pela falta “crônica” de recursos — é levada em consideração nessa segunda edição do livro.

A preservação digital deve começar pelos conceitos básicos de arquivamento, do qual a preservação é apenas uma parte. O trabalho de arquivamento envolve atividades como avaliação, organização, descrição (catalogação), preservação, gerenciamento e acesso.¹⁹⁷

¹⁹² Dilema Digital 2. 2015, p.1.

¹⁹³ Dilema Digital 1. 2009, p. 1.

¹⁹⁴ Dilema Digital 1. 2009, p. 1.

¹⁹⁵ Dilema Digital 1. 2009, p.1.

¹⁹⁶ Dilema Digital 2. 2015, p.1.

¹⁹⁷ Dilema Digital 2. 2015, p.42.

O livro traz, ainda, algumas definições vigentes acerca da Preservação Digital:

A preservação digital combina políticas, estratégias e ações para garantir o acesso a materiais nativos digitais, ou convertidos em formatos digitais, independentemente dos riscos de falhas na mídia ou nas mudanças tecnológicas. O objetivo da preservação digital é a reconstituição precisa de conteúdo tido como autenticado, ao longo do tempo.

– Associação de Acervos de Bibliotecas e Serviços Técnicos

A preservação digital consiste em uma série de ações e intervenções necessárias para garantir o acesso confiável e contínuo de objetos digitais autênticos, pelo tempo que forem considerados valiosos. Essas ações e intervenções abrangem não apenas as atividades técnicas, mas também todas as considerações estratégicas e organizacionais relativas à sobrevivência e ao gerenciamento do material digital.

- Comitê Conjunto de Sistemas de Informação do Reino Unido

Objetos digitais deixarão de ser acessíveis sem o gerenciamento e intervenção ativos. O maior risco ao acesso a objetos digitais é o desenvolvimento contínuo de hardware e software. Muitos formatos ou arquivos digitais dependem de um ambiente computacional particular para uma representação precisa de seu conteúdo. Qualquer modificação no ambiente de renderização dos dados pode resultar em modificações na representação da reconstituição de um determinado recurso – ou resultar na total impossibilidade de sua reconstituição.

- Idem ¹⁹⁸

Para os autores do livro, o termo “preservação digital” não é muito preciso, preferindo o termo "arquivamento digital", que é diferente da compreensão trazida por profissionais de tecnologia da informação tradicionais para o termo, em que o termo “arquivamento” diz respeito à realização regular de cópias de segurança (*backup*) em fitas, com ou sem fases de verificação. Outros preferem o termo "curadoria de dados", que destaca a gestão ativa dos arquivos digitais durante sua existência.¹⁹⁹

Independentemente do termo utilizado, a preservação de mídias digitais é um processo ativo que abrange todos os estágios do trabalho de arquivamento tradicional e outras ações adicionais importantes, algumas das quais de alta complexidade. Algumas dessas ações adicionais devem ser tomadas antes mesmo da criação dos dados audiovisuais digitais – a seleção dos formatos de arquivos digitais e da mídia de armazenamento, por exemplo, e a avaliação e escolha de fluxos de trabalho que facilitem o processo de preservação.²⁰⁰

¹⁹⁸ Dilema Digital 2. 2015, p.43.

¹⁹⁹ www.cc.ac.uk/about-us/dcc-charter apud Dilema Digital 2. 2015, p.43.

²⁰⁰ Dilema Digital 2. 2015, p. 43.

Dessa forma, preservar arquivos digitais constitui um processo que demanda também outras práticas como “a copiagem para novas mídias e formatos de arquivos de dados (‘migração’), a manutenção de redundância (que significa verificar se não há dados repetidos) e verificação da integridade dos dados”,²⁰¹ além do planejamento e gestão. Para tanto, deve-se levar em consideração que “a avaliação e a catalogação de acervos digitais demandam ferramentas de hardware e software especializadas, que podem variar de acervo para acervo”²⁰² e o trabalho com tais ferramentas demanda uma mão de obra especializada.

O autor ainda aponta a importância de os “arquivos de dados” serem acomodados em algum lugar, “preferencialmente em um ambiente no qual o sistema de armazenamento facilite as ações de preservação digital e mesmo a automação de muitas delas”²⁰³, o que é bastante improvável no Brasil. Os sistemas atuais de armazenamento digital podem variar em versões simples e portáteis, como os HDs externos comprados em lojas comuns, ou altamente complexos como um centro de processamento de dados (*data center*) incorporado à infraestrutura de TI da instituição. “Esses sistemas maiores, e geralmente gerenciados ativamente, também são chamados de ‘sistemas de armazenamento centralizado’ (*enterprise storage systems*)”.²⁰⁴

Importante ressaltar que, independentemente do sistema de armazenamento utilizado, o acesso é usualmente feito por meio de um computador com as conexões de hardware, software e de rede (internet) apropriadas. E mais relevante, ainda, é compreender que “o armazenamento digital em si não é suficiente para efeitos de preservação”²⁰⁵. Se faz necessário, então, que ações próprias da preservação sejam executadas e supervisionadas.

²⁰¹ Dilema Digital 2. 2015, p. 44.

²⁰² Dilema Digital 2. 2015, p. 44.

²⁰³ Dilema Digital 2. 2015, p. 46.

²⁰⁴ Dilema Digital 2. 2015, p. 46.

²⁰⁵ Dilema Digital 2. 2015, p. 46.

O monitoramento pode ser realizado utilizando softwares simples de gerenciamento de base de dados como FileMaker Pro ou Microsoft Access, ou um software completo conhecido como sistema de gerenciamento de ativos digitais (DAM ou DAMS - Digital Asset Management). A implementação efetiva de ações de preservação também implica na utilização de outras ferramentas de software, tais como leitores de arquivos digitais ou transcodificadores (que convertem um formato de arquivo digital para outro).

Segundo os autores, existem benefícios e contrapartidas a serem levados em consideração quanto à escolha dos softwares para fins arquivísticos. Tal assunto foi aprofundado nos estudos de caso presentes no Dilema Digital, denominado "*Long-Term Management and Storage of Digital Motion Picture Materials* (Gerenciamento e armazenamento a longo prazo de materiais cinematográficos digitais)"²⁰⁶.

A seguir, a compreensão das práticas básicas de preservação que são, a saber: cópias de segurança dos dados (*backup*), verificação e migração.

A preservação de arquivos digitais é um esforço ativo e permanente, na qual a probabilidade da manutenção do acesso aos dados aumenta se forem realizadas três atividades [acima citadas e abaixo descritas]. A falha na implementação prática de qualquer uma dessas atividades básicas poderá, eventualmente, resultar na perda de dados.²⁰⁷

Cópias de segurança de dados (*backup*): "realização de múltiplas (duas ou mais) cópias de um arquivo digital. As cópias devem ser armazenadas em diferentes locais geográficos e em diferentes tipos de mídia para proteção contra desastres físicos ou técnicos".

Verificação: "inspeção de rotina de todas as cópias de arquivos de dados para proteção contra falhas na mídia ou na transferência dos dados. Uma atividade relacionada é a verificação da integridade (*fixity checking*), que atesta se o arquivo digital foi modificado, com ou sem intenção".

Migração: "transferência regular de todas as cópias dos arquivos digitais para mídias e formatos de arquivos mais recentes para proteção contra a obsolescência tecnológica".²⁰⁸

²⁰⁶ Dilema Digital 2. 2015, p. 46.

²⁰⁷ Dilema Digital 2. 2015, p. 46.

²⁰⁸ Digital Dilema 2. 2015, p. 46.

4.1 Dossiê de ações

As ferramentas a seguir são opções possíveis para a preservação audiovisual. A eficácia de cada uma delas dependerá do contexto de cada instituição ou usuário, então, antes mesmo de optar por uma ou outra, se faz necessário montar um fluxo de trabalho, considerando-se o tamanho do acervo, contexto econômico e profissionais disponíveis. O plano de preservação de cada lugar deve ser montado com base em uma série de critérios e desenvolvimento posterior das ações de salvaguarda e manutenção dos acervos.

Para melhor entendimento desse fluxo de trabalho, os guias práticos²⁰⁹ e também o glossário da DPC online (*Digital Preservation Coalition*²¹⁰ ou Coligação de Preservação Digital) encontram-se nos anexos deste trabalho. Sucintos e ilustrados, ainda não foram traduzidos para o português, mas trazem grandes *insights* para a implementação de meios de preservação digital.

Uma grande iniciativa para ajudar pessoas e instituições é o COPTR (*Community Owned digital Preservation Tool Registry* ou Comunidade detentora de ferramentas de registro de preservação digital).²¹¹ A comunidade tem por função agrupar e descrever as ferramentas que são úteis no processo da preservação digital em longo prazo e funciona também como mecanismo de descoberta e avaliação, que auxilia os profissionais em busca de soluções para preservar os dados digitais. O COPTR pretende reunir, em um só lugar, o conhecimento da comunidade de preservação digital. O grupo captura detalhes básicos sobre uma ferramenta: o que ela faz, como encontrar mais informações (URLs relevantes) e referências às experiências dos usuários com a ferramenta.²¹²

Um outro empreendimento na área é o APARSEN (*Alliance for Permanent Access Network*) tools registry²¹³, que funciona como um repositório que tenta construir uma base de evidências para ferramentas de preservação e, em particular, busca

²⁰⁹ DPC ONLINE. Manuais de preservação. Disponíveis em: <https://dpconline.org/handbook> e <https://dpconline.org/handbook/glossary> Ativo em 01 mar 2018. Disponível nos anexos.

²¹⁰ <https://www.dpconline.org/> ativo em 01 mar 2018.

²¹¹ http://coptr.digipres.org/Main_Page

²¹² Informações retiradas da cartilha Tech solutions and tools, do DPC group. Tradução Livre.

²¹³ <http://www.alliancepermanentaccess.org/index.php/tools/tools-for-preservation/>

identificar quais ferramentas são apropriadas para o tipo de dados, inclusive, os audiovisuais ²¹⁴. A APARSEN recolhe detalhes de software relacionado à preservação, exemplos de dados e evidências de software de ligação de preservação para tipos de dados. Algumas dessas evidências provêm de testes específicos, mas muito vem dos cenários dos usuários. O recurso é agora mantido *pela Alliance for Permanent Access* (ou Aliança para Acesso Permanente – APA)²¹⁵.

A *AV Preserve tools list*²¹⁶ (ou Lista de ferramentas de preservação AV) é uma lista de ferramentas de uso particular na preservação em longo prazo de materiais audiovisuais, tanto digitalizados como nativos digitais²¹⁷.

O *Digital Curation Centre (DCC) tools and services list*²¹⁸ (ou Centro de curadoria digital) é um centro que apoia pesquisadores no Reino Unido no enfrentamento dos desafios para a preservação de recursos digitais. Para tanto, oferece uma série de serviços de suporte e assessoria diretamente ligados à pesquisa e desenvolvimento. Já fizeram um catálogo de ferramentas e serviços pesquisadores e curadores. Tal informação também está integrada ao COPTR, que citamos anteriormente²¹⁹.

O *Digital Cultural Heritage Roadmap for Preservation*²²⁰ (ou Mapa Patrimonial Cultural Digital para Preservação – DCH-RP) funciona como um roteiro, cuja iniciativa coletou e descreveu informações e conhecimentos relacionados a ferramentas, tecnologias e sistemas que podem ser aplicados para fins de preservação do patrimônio cultural digital. Sua mais recente versão foi criada em 2014²²¹.

O *Inventory of FLOSS*²²² – *Free/libre open-source software* (ou Inventário com software de código-aberto/ livre para uso) foi produzido e financiado por instituições

²¹⁴ Informações retiradas da cartilha Tech solutions and tools, do DPC group. Tradução livre.

²¹⁵ Informações retiradas da cartilha Tech solutions and tools, do DPC group. Tradução livre.

²¹⁶ <http://www.avpreserve.com/avpsresources/tools/>

²¹⁷ Informações retiradas da cartilha Tech solutions and tools, do DPC group. Tradução livre.

²¹⁸ <http://www.dcc.ac.uk/resources/external/tools-services>

²¹⁹ Informações retiradas da cartilha Tech solutions and tools, do DPC group. Tradução livre.

²²⁰ <http://www.dch-rp.eu/index.php?en/137/registry-of-services-tools>

²²¹ Informações retiradas da cartilha Tech solutions and tools, do DPC group. Tradução livre.

²²² https://docs.google.com/spreadsheet/ccc?key=0Ag_7rVJwt0CpdFRJOEJxdEk4ZEMxQ01jaDgxQXFSTkE#gid=0

européias, e produziu uma lista de softwares de código aberto e gratuitos que podem ser úteis no setor de preservação do patrimônio cultural. Embora não esteja limitado às ferramentas de preservação digital, contém informações sobre uma grande variedade de ferramentas com aplicativos de preservação digital, avaliando seu propósito, qualidade de documentação, nível de suporte, requisitos de licença e fornecendo links para informações do projeto e código-fonte²²³.

Essas são apenas algumas iniciativas que procuram mapear e listar as ferramentas possíveis, bem como avaliar o funcionamento de cada uma delas. Considerando as possibilidades de armazenamento em nuvem, os provedores de serviços abaixo podem ser bastante úteis no processo.

O *ArchivesDirect*²²⁴ (ou arquivo direto) é um serviço para a criação de pacotes de conteúdo de preservação digital baseados em padrões arquivados em armazenamento seguro a longo prazo²²⁵. Ele funciona como uma instância hospedada de Archivemata (que explicaremos adiante) com armazenamento seguro em nuvem, replicado em diferentes servidores²²⁶.

O *Arkivum*²²⁷ é uma iniciativa de serviço totalmente gerenciado e seguro para a retenção de dados de longo prazo com acesso *online* e uma garantia de integridade de dados que é parte do seu Acordo de Nível de Serviço e respaldado por seguros mundiais²²⁸.

Já o *DuraCloud*²²⁹ é um serviço gerenciado da *DuraSpace*²³⁰ (uma iniciativa cujo objetivo é preservar e proporcionar acesso ao patrimônio intelectual, cultural e científico do mundo). A ferramenta fornece suporte e opções que copiam, automaticamente, conteúdo para vários provedores de armazenamento em nuvem

²²³ Informações retiradas da cartilha Tech solutions and tools, do DPC group. Tradução livre.

²²⁴ <http://archivesdirect.org>

²²⁵ Informações retiradas do site <http://archivesdirect.org>. Tradução livre.

²²⁶ Informações retiradas da cartilha Tech solutions and tools, do DPC group. Tradução livre.

²²⁷ <http://arkivum.com>

²²⁸ Informações retiradas da cartilha Tech solutions and tools, do DPC group. Tradução livre.

²²⁹ <http://www.duracloud.org>

²³⁰ <http://www.duraspace.org/> é uma iniciativa

diferentes e garantem que todas as cópias do conteúdo permaneçam sincronizadas²³¹.

O *Preservica*²³² é uma solução de armazenamento em nuvem para preservação ativa dos arquivos de pequenas e médias organizações, ou mesmo acervos particulares. A versão para nuvem, denominada *Preservica Cloud Edition*, é uma plataforma de preservação digital compatível com o esquema OAIS, totalmente equipada com nuvem, que também inclui acesso/pesquisa do público para permitir que você compartilhe com segurança seu arquivo ou coleção²³³.

A *Plato*²³⁴ é uma ferramenta de monitoramento e planejamento de preservação para organizações encarregadas de proteger materiais digitais. “O software orienta os usuários através de um processo de definição de requisitos, avaliação de processos potenciais e uso dos resultados para facilitar a tomada de decisões na criação de um plano de preservação final”²³⁵ e sugere formatos e ações para seus arquivos.

O LOCKSS ²³⁶ (*Lots Of Copies Keep Stuff Safe* – ou Muitas cópias mantêm as coisas seguras) é um programa oriundo das Bibliotecas da Universidade de Stanford, nos EUA, que oferece “bibliotecas e editores com ferramentas de preservação digital de código aberto premiadas e de baixo custo para preservar e fornecer acesso a conteúdos digitais persistentes e autoritários”²³⁷.

A COPPUL²³⁸ (Council of Prairie and Pacific University Libraries) é uma rede de preservação digital que cria capacidade para a preservação digital nas instituições membros, “oferecendo infraestrutura de preservação, oportunidades de educação e recursos de advocacia, conectando atividades locais de preservação digital com esforços regionais, nacionais e internacionais”²³⁹.

²³¹ Informações retiradas da cartilha Tech solutions and tools, do DPC group. Tradução livre.

²³² <http://preservica.com/edition/cloud-edition/>

²³³ Informações retiradas da cartilha Tech solutions and tools, do DPC group. Tradução livre.

²³⁴ <http://www.ifs.tuwien.ac.at/dp/plato/intro/> ativo em 05 mar 2018

²³⁵ <http://www.dcc.ac.uk/resources/external/plato> tradução livre ativo em 05 mar 2018

²³⁶ <https://www.lockss.org/>

²³⁷ <https://www.lockss.org/> tradução livre ativo em 05 mar 2018

²³⁸ <https://coppul.ca/>

²³⁹ <https://coppul.ca/digital-preservation> tradução livre ativo em 05 mar 2018

O PRONOM²⁴⁰ é um projeto executado e mantido pelo Arquivo Nacional do Reino Unido. Esse recurso funciona com o registro *online* de informações técnicas. Eficiente para resoluções imparciais e definitivas sobre os formatos de arquivos digitais, “produtos de software e outros componentes técnicos necessários para suportar o acesso a longo prazo a registros eletrônicos e outros objetos digitais de valor cultural, histórico ou comercial”²⁴¹.

Atualmente também existem empresas que oferecem o armazenamento e preservação de arquivos digitais. Esse tipo de serviço é a solução encontrada por quem se preocupa com seus arquivos, mas não sabe como proceder e pode pagar por isso. Isso vale para arquivos digitais de qualquer natureza, abaixo apresentaremos, apenas, as iniciativas que contemplam os arquivos audiovisuais.

Fora do Brasil, temos um renomado instituto com esse fim, o INA France – *Institut national de l'audiovisuel*²⁴², criado em 1975 como um repositório dos arquivos de rádio e TV da França²⁴³. A instituição partia do pressuposto de que “todas as organizações com arquivos ou produção de conteúdo [enfrentavam] os mesmos desafios estratégicos: conservar, digitalizar e valorizar seus fundos”²⁴⁴ e, portanto, eles fariam um trabalho de apoio a essas organizações. Esse apoio se dá das mais diversas formas: “conhecimentos de fundo, recomendações técnicas, assistência de gerenciamento de projetos, conselhos estratégicos e treinamento em equipe”²⁴⁵.

A equipe de trabalho conta com especialistas de diversas áreas para atender as demandas e projetos que chegam, da França e do exterior. Em poucas palavras, o INA é “uma instituição precursora para digitalização e aprimoramento de arquivos audiovisuais, (...) e desenvolveu habilidades transversais [que permitiram que ele se tornasse] uma referência para arquivos e mídia”.

Abaixo, a tradução do esquema de trabalho da instituição:

²⁴⁰ <http://www.nationalarchives.gov.uk/PRONOM/Default.aspx>

²⁴¹ <http://www.nationalarchives.gov.uk/aboutapps/PRONOM/tools.htm>. Tradução livre. Ativo em 05 mar 2018

²⁴² Instituto Nacional de Audiovisual.

²⁴³ Palestra com Benjamin Lerena – Gerente de Projeto, Especialista em Gestão de Ativos Audiovisuais do INA, na 11ª edição do Cine OP na data XX de 2017.

²⁴⁴ <https://www.ina-expert.com/expertise-conseil-et-etudes> tradução livre Acesso em: 12 fev. 2018.

²⁴⁵ Idem, tradução livre.



FIGURA 35 – Fluxo de trabalho do INA

Fonte: <https://www.ina-expert.com/expertise-conseil-et-etudes/metiers.html>

Tradução livre. Acesso em 18 fev. 2018

As frentes de trabalho consideram o fluxo acima, tendo como principal objetivo a valorização dos conteúdos, que se dá por meio da digitalização e preservação das fontes, bem como gestão e documentação do acervo.

Para eles, a digitalização “é um meio de salvaguardar e transmitir o seu patrimônio, mas a transição da tecnologia analógica para a tecnologia digital conduz a uma interrupção das práticas”. Para tanto, se faz necessária a presença de especialistas no assunto, capazes de planejar e executar a digitalização, sendo feita principalmente em mídias analógicas, como “fitas de áudio e vídeo (1 polegada, 2 polegadas), cassetes (Beta, U-matic, ...), filmes (16mm, 35mm, ...) e fotos”²⁴⁶.

Para a documentação, a visão do INA é de que “o acesso a documentos audiovisuais envolve a descrição do conteúdo e a gestão da informação. O design de sistemas de documentos adaptados à natureza e uso dos arquivos”²⁴⁷ é prioridade nos projetos da instituição, bem como otimização das ferramentas que

²⁴⁶ Site <https://www.ina-expert.com/expertise-conseil-et-etudes/metiers.html> tradução livre

²⁴⁷ Idem.

serão utilizadas e consultoria personalizada de acordo com a realidade de cada acervo. As tecnologias utilizadas, para escanear e digitalizar os acervos, são avançadas e a qualidade é incontestável, além da garantia de salvaguarda dos arquivos.

Além do INA, existem também outras empresas e iniciativas nesse sentido em algumas partes do mundo. No Brasil não temos, até o momento, nenhuma empresa especializada especificamente em acervos audiovisuais, mas algumas opções para gestão estratégica de documentos e segurança da informação de um modo geral. Este trabalho não se preocupa em citá-las porque a estrutura utilizada não se adequa às necessidades do audiovisual, e pagar por esse tipo de serviço ainda não é algo factível à realidade brasileira dos arquivos e acervos.

E temos, também, as iniciativas públicas, que, atualmente, se baseiam na construção de repositórios *online* e em canais na internet onde é possível guardar para acesso.

Uma opção que vem, com frequência, sendo utilizada no Brasil é o software livre “archivematica”:

Archivematica é um sistema gratuito de preservação digital de código aberto, que possibilita o acesso de longo prazo da memória digital. O conteúdo preservado no Archivematica pode ser acessado via AtoM, um sistema de gerenciamento de conteúdo Web. Desenvolvido pela empresa canadense Artefactual Systems, com a colaboração de instituições como a UNESCO, por meio do *Memory of the World's - Subcommittee on Technology*, o Arquivo Municipal da cidade de Vancouver, o Arquivo Central de Rockefeller, o Arquivo da Universidade Simon Fraser University, a Biblioteca da Universidade de British Columbia e outros. Archivematica está sob AGPL e licenças de código aberto *Creative Commons*, e sua estrutura de preservação digital está em conformidade com o modelo ISO-OAIS e outras normas de preservação digital²⁴⁸.

²⁴⁸ Infos retiradas do guia do usuário: http://wiki.ibict.br/index.php/Guia_do_Usu%C3%A1rio_-_Archivematica acesso em 18 fev. 2018.



FIGURA 36 – página inicial do Archivemática.
 Fonte: <https://www.archivemática.org/en/> Acesso em 20 fev 2018

Repositórios como esse têm por função utilizar “estratégias de preservação de longo prazo, evitando a obsolescência tecnológica, a incompatibilidade de formato, versões e suportes”²⁴⁹ bem como a garantia de que os arquivos digitais permanecerão “acessíveis e utilizáveis ao longo dos anos”²⁵⁰.

O método de preservação que o Archivemática propõe é o de normalização dos arquivos utilizando uma prática denominada *ingest*. *Ingest* é, segundo o Manual de comunicação brasileiro:

1 – Inserção de material audiovisual nos servidores de armazenamento ou ilhas de edição não linear [no caso das TVs]; e 2 – Processo pelo qual um arquivo digital é aceito e carregado em um sistema de armazenamento digital, junto com o cadastro correspondente para posterior recuperação [no caso dos arquivos].²⁵¹

Em seguida, o sistema cria cópias de preservação e cópias de acesso. “Note que os arquivos originais são mantidos sempre, para permitir diferentes ações de preservação no futuro, tais como a normalização de diferentes formatos de arquivo

²⁴⁹ FONTANA, F. F. et al. 2014, p.11.

²⁵⁰ FONTANA, F. F. et al. 2014, p.11.

²⁵¹ <https://www12.senado.leg.br/manualdecomunicacao/glossario/ingest> acesso em 18 fev. 2018.

ou emulação”²⁵². O sistema é atualmente recomendado pelo Arquivo Nacional do Brasil para as práticas de preservação de acervos.²⁵³

O archivemática utiliza a estrutura OAIS, já citada neste trabalho, que significa *Open Archival Information System*. Tal modelo é padronizado pela ISO²⁵⁴ e foi traduzido e adaptado para o Brasil como SAAI – Sistema Aberto de Arquivamento de Informações, objetivando “a orientação de um sistema de arquivo dedicado à preservação e manutenção do acesso a informações digitais em longo prazo”²⁵⁵.

O modelo é bastante funcional por considerar as três instâncias de contato com o acervo, a saber: produtores, administradores e consumidores.



FIGURA 37 – Tradução da versão simplificada do modelo OAIS. Encontrada em: <https://www.slideshare.net/redecariniana/preservao-digital-de-arquivos-audiovisuais> disponível em 23 fev 2018.

²⁵² Infos retiradas do guia do usuário: http://wiki.ibict.br/index.php/Guia_do_Usu%C3%A1rio_-_Archivem%C3%A1tica acesso em 18 fev. 2018.

²⁵³ FONTANA, F. F. et al. 2014, p.10.

²⁵⁴ FONTANA, F. F. et al. 2014, p.10.

²⁵⁵ FONTANA, F. F. et al. 2014, p.10.

A função de cada uma dessas instâncias está descrita a seguir:

PRODUTOR: pessoas ou sistemas que fornecem a informação a ser submetida e preservada via pacotes de informação;

ADMINISTRADOR: entidade responsável pelo estabelecimento de políticas e regras gerais do arquivo, como a escolha do formato a ser adotado como padrão para acesso e preservação de longo prazo;

CONSUMIDOR: pessoas ou sistemas que irão interagir com os serviços OAI seja através de um repositório ou sistema de gerenciamento arquivístico. Cabe ressaltar que os consumidores são aqueles que acessam as informações e que são capazes de compreender a informação preservada, ou seja, é uma comunidade-alvo específica.²⁵⁶

“O modelo OAI, através de seus micro-serviços visa preservar e dar acesso a informações de interesse para que estejam acessíveis a longo prazo, num sistema que garanta a autenticidade e integridade do objeto digital”²⁵⁷.

No Brasil está em andamento uma grande pesquisa com um propósito semelhante ao que o Archivematica faz, pesquisadores investem na criação de um projeto denominado Legatum²⁵⁸: “um repositório cooperativo, coletivo, multilíngue, reunindo acervos de instituições públicas, que visa à preservação digital e ao acesso remoto a acervos audiovisuais de diferentes instituições, num mesmo ambiente digital, em versão beta, em constante aperfeiçoamento e evolução, com conclusão prevista para fevereiro de 2019” da Universidade Federal da Bahia²⁵⁹ com propósito de criar “um modelo arquivístico de preservação digital do patrimônio audiovisual custodiado por instituições brasileiras e de outros países de idiomas de origem latina”.

Um artigo²⁶⁰ dos pesquisadores do projeto apresenta o percurso e também os resultados provisórios da pesquisa. A proposta é desenvolver

um modelo de repositório destinado ao acesso remoto a versões digitais nativas ou resultantes de processos de conversão de itens constituintes de acervos audiovisuais públicos de instituições brasileiras e de outros países de idiomas de origem latina. A iniciativa, denominada Legatum, constitui-se em uma série de ações coordenadas ao redor de um repositório aberto,

²⁵⁶ FONTANA, F. F. et al. 2014, p.10.

²⁵⁷ FONTANA, F. F. et al. 2014, p.10.

²⁵⁸ <http://www.legatum-si.net/atom/>

²⁵⁹ <http://www.legatum.ufba.br/web/>

²⁶⁰ SILVA, R. D. et al. 2016.

multi-idiomas e colaborativo, baseada em uma plataforma digital compatível com normas arquivísticas internacionais²⁶¹.

Tal trabalho está sendo desenvolvido no grupo CRIDI (Grupo de Estudos sobre Cultura, Representação e Informação Digitais) da Universidade Federal da Bahia, desde 2005. A pesquisa foi dividida em fases, e inicialmente coletou “dados sobre acervos públicos fotográficos, sonoros e audiovisuais em 109 instituições públicas na cidade do Salvador, sendo que entre essas instituições, somente 35 mantinham documentação audiovisual”²⁶².

O grupo ainda se dedica à participação de eventos na área e iniciou a oferta de disciplinas relacionadas na própria universidade. Posteriormente, decidiram limitar o trabalho, apenas, aos acervos audiovisuais, refazendo os questionários e fazendo análises comparativas que possibilitam conhecer “a evolução, involução ou estagnação da preservação destas instituições”²⁶³. Dessa forma, o grupo pretendia “inferir e avaliar o nível de compromisso das esferas públicas decisórias acerca da salvaguarda do patrimônio audiovisual”²⁶⁴.

Os questionários aplicados eram bastante completos e trouxeram ao grupo informações valiosas que identificavam, a seguir:

1) temas que caracterizam os acervos; 2) formatos e suportes existentes; 3) instrumentos utilizados para o controle de consultas ao acervo; 4) serviços e equipamentos oferecidos aos usuários; 5) indicações sobre prioridades para uma primeira ação de conversão digital; 6) o conhecimento técnico das equipes, relativas a tratamento de acervos originais (analógicos) e de digitalização de acervos de imagem em movimento; 7) equipamentos disponíveis nos setores de documentação audiovisual; 8) existência de normativas internas, políticas de preservação e planos de desastres para a salvaguarda das imagens da instituição; 9) experiências prévias de digitalização de acervos de imagem em movimento, bem como sobre a constituição das equipes e eventuais consultorias contratadas para esta finalidade; 10) a origem dos recursos financeiros e materiais de tais experiências; 11) existência de relatórios técnicos de experiências eventualmente realizadas de conversão digital de imagem em movimento; 12) aspectos predominantes dos processos decisórios sobre o que digitalizar em cada instituição, bem como se as operações de digitalização foram realizadas internamente ou se foram contratadas externamente; 13) aspectos técnicos do processo de digitalização.²⁶⁵

²⁶¹ SILVA, R. D. et al. 2016, p. 1.

²⁶² SILVA, R. D. et al. 2016, p. 5.

²⁶³ SILVA, R. D. et al. 2016, p. 6.

²⁶⁴ SILVA, R. D. et al. 2016, p. 6.

²⁶⁵ SILVA, R. D. et al. 2016, p. 6 -7.

Dentre os objetivos dessa coleta, o grupo pretendia entender “o quê, como e porquê vem sendo feito em termos de digitalização de acervos audiovisuais analógicos e identificar informação relevante para extrair dados destas experiências”²⁶⁶ e, portanto, reunir dados que pudessem ser úteis para as ações que irão propor.

Antes do final do século passado, a preservação significava reter, guardar. Atualmente ela passa a se constituir como uma área associada ao acesso, formando um binômio em que preservar significa compartilhar, difundir. No universo digital, preservação e salvaguarda passam a ser sinônimos de acesso continuado, pelo maior tempo possível. Assim, prover meios de acesso ao legado registrado nos documentos de valor permanente e identificados como pertencentes à sociedade nos quais foram produzidos é um fator determinante do momento atual, que extrapola o paradigma da restrição ao manuseio do original quando a preservação era o grande objetivo.²⁶⁷

O repositório Legatum funciona como uma plataforma cujo acesso se dá de forma digital. Semelhante ao archivematica, o usuário precisa instalar o software livre AtoM. O Legatum que pode ser acessado pelo site <http://www.legatum-si.net> é ainda um modelo. Sua pretensão de funcionamento é se estabelecer como um “repositório aberto destinado ao acesso digital de informações (...) e, com potencial para se transformar em um repositório arquivístico confiável para armazenamento de matrizes digitais audiovisuais”²⁶⁸.

A imagem a seguir (FIGURA 38) revela o campo de busca de instituições que, atualmente, apoiam ou já possuem acervos na versão beta.

A preocupação dos realizadores é de que o repositório funcione “em conformidade com normas arquivísticas, de forma a se tornar tão adequado quanto possível para alcançar os objetivos a que se propõe”²⁶⁹.

²⁶⁶ SILVA, R. D. et al. 2016, p. 7.

²⁶⁷ SILVA, R. D. et al. 2016, p. 8.

²⁶⁸ SILVA, R. D. et al. 2016, p. 20.

²⁶⁹ SILVA, R. D. et al. 2016, p. 20.

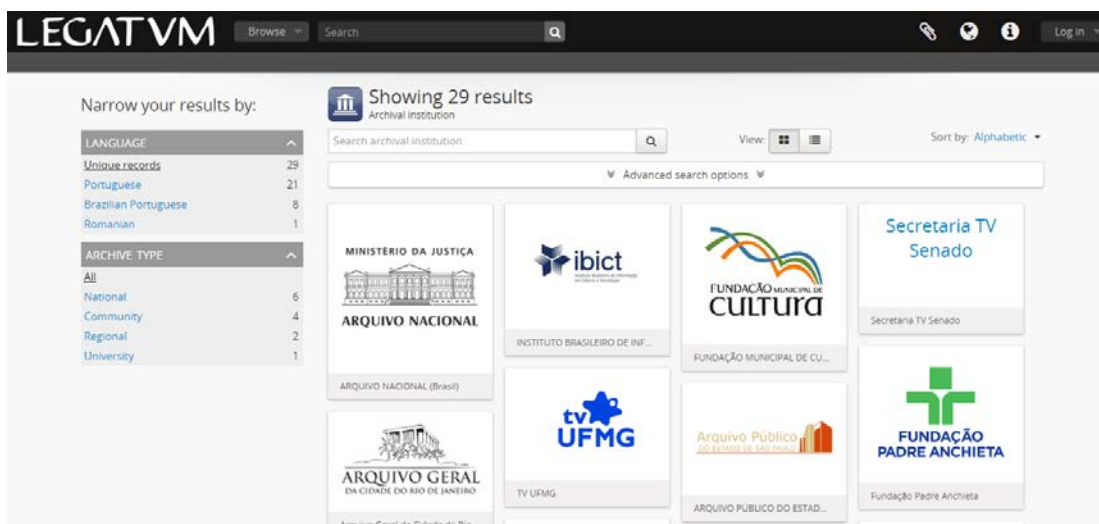


FIGURA 38 – Captura de uma das telas online do Sistema Legatum.
 Fonte: <http://www.legatum.ufba.br/atom/index.php/repository/browse>
 Ativo em 22 fev 2018.

Legatum, com o apoio institucional da esfera pública, pode proporcionar uma solução de custo tecnológico razoável e de baixo risco (embora requeira muito trabalho de convencimento junto às instituições que queremos ver utilizando o ambiente), ao mesmo tempo em que oferece elementos de reflexão teórica e prática operacional, reunindo esclarecimentos técnicos e até mesmo, acreditamos, estímulos ao engajamento também dos futuros usuários finais do ambiente.²⁷⁰

Os organizadores do Legatum relatam ainda encontrarem dificuldade no diálogo e na sensibilização das instituições, motivo pelo qual este trabalho também se faz relevante. Segundo os pesquisadores, “talvez esta seja uma das maiores dificuldades [que estão] enfrentando”²⁷¹.

O desafio está posto, de certa forma, ao trabalharmos na interface arquivística da ciência da informação: vamos procurar novas trilhas, procurando argumentar no desenvolvimento do projeto sobre a importância do acesso digital aos documentos audiovisuais públicos como instância de ampliação da consciência, de geração de novas vozes que fazem parte do processo histórico, dos estudantes, dos educadores, dos curiosos, deste mundo de subjetividade que a internet possibilita.²⁷²

O MIS BH, antigo CRAV em Belo Horizonte, começou, em 2015 um projeto que colocava num canal do YouTube alguns vídeos da instituição. Os filmes que utilizam no Canal MIS²⁷³ são produções de órgãos da Prefeitura de Belo Horizonte e filmes realizados pelo CRAV/MIS ou fomentados por algum edital da própria instituição,

²⁷⁰ SILVA, R. D. et al. 2016, p 23.

²⁷¹ SILVA, R. D. et al. 2016, p 23.

²⁷² SILVA, R. D. et al. 2016, p 10.

²⁷³ <https://www.youtube.com/CanalCRAV>

portanto, filmes cujo direito de divulgação pertence ao museu. O canal é alimentado semanalmente. Uma iniciativa válida para conhecimento do acervo pelo público, bem como ação paliativa de preservação. Afinal, já foi dito neste trabalho que o armazenamento digital em si não é suficiente para efeitos de preservação e o armazenamento na nuvem (mesmo possuindo o arquivo salvo em outros lugares e formatos) ainda não é considerado confiável. Portanto, somente ter vídeos em um canal do YouTube garantiria, apenas, o acesso prático de uma pesquisa, por exemplo.

O potencial do YouTube como um arquivo audiovisual pode ser descrito somente após as principais características que definem o arquivo tradicional serem apresentadas.

Governado por proveniência, custódia e uma autoridade central, um arquivo tradicional pode ser, geralmente, considerado como um conjunto de registros que tem um significado histórico, cultural ou de prova. O princípio de arquivo de proveniência, que lida com a origem de registros, tem significados distintos. O uso mais comum refere-se ao "escritório de origem" de registros ou para a entidade administrativa — individual, familiar ou instituição — a partir do qual os registros, documentos pessoais ou manuscritos derivam.²⁷⁴

YouTube, segundo a professora e pesquisadora Sheenagh Pietrobruno²⁷⁵, nunca pode ser um arquivo, se a versão tradicional e o modelo que esse serviço de hospedagem de vídeo propõe é emular. Os arquivos do patrimônio imaterial no YouTube, que estão continuamente sendo construídos pelos usuários, não possuem um claro ponto de origem ou um conjunto de documentos iniciais. As crescentes imagens em movimento do YouTube são produzidas pelos próprios usuários, como vimos no capítulo 3 deste trabalho. O que é considerado valioso e digno de guarda é até certo ponto determinado por usuários, em vez de por uma autoridade central e, nesse caso, os usuários do YouTube não têm controle absoluto, como o Google

²⁷⁴ PIETROBRUNO, Sheenagh. 2013, p.3.

²⁷⁵ <http://www.sheenaghpietrobruno.com/>

mantém o direito de remover qualquer vídeo cujo conteúdo considere inapropriado²⁷⁶.

No entanto, uma grande parte do conteúdo no YouTube é criada e enviada pelos usuários. Numerosos vídeos gerados por usuários carregados no YouTube poderiam ser considerados como fontes primárias, como são divulgados primeiro sobre este serviço de hospedagem de vídeo — no entanto, bem sabemos que muitos vídeos do YouTube são derivados de mídia convencional e propaganda comercial²⁷⁷.

O MIS BH, também, trabalhou alguns pontos desde o começo dessa pesquisa. Em 2017, o museu fez o levantamento da quantidade de espaço em disco rígido e de tamanho dos arquivos existentes no acervo e também recebeu mais alguns HDs externos com filmes resultantes de um edital de Cultura Popular.

Em 2017 também foi continuado o trabalho de organização documental no MIS. Foram discutidas as necessidades do museu em relação à organização e ao arranjo dos documentos produzidos e daqueles que já fazem parte do acervo da instituição. Após a definição dos parâmetros, foi redigido o Plano de Classificação de Documentos, que orienta sobre a maneira de se classificar e arquivar documentos digitais e impressos no MIS. Após a finalização do documento, foram realizadas reuniões com o objetivo de orientar toda equipe sobre os procedimentos referentes aos documentos gerados na instituição²⁷⁸.

Tal resolução definiu que todos os funcionários trabalhariam os documentos sob sua responsabilidade seguindo as orientações do Plano de Classificação e que todos os novos documentos gerados seriam, sempre que possível, inseridos já na nova organização digital dos arquivos compartilhados no museu.

²⁷⁶ PIETROBRUNO, Sheenagh. 2013, p.3.

²⁷⁷ PIETROBRUNO, Sheenagh. 2013, p.3.

²⁷⁸ Informações obtidas em conversa informal com Marcella do CRAV.

Já em 2018, os funcionários estão trabalhando na elaboração de um manual sobre o digital específico para o MIS, a expectativa é de que ele contemple normas de digitalização e preservação digital que se encaixem na realidade do museu.

Outra iniciativa é o Banco de Conteúdos Culturais – BCC ²⁷⁹, que trabalha com a digitalização de acervos e banco de conteúdos audiovisuais brasileiros e implantação de política nacional de digitalização e democratização do acesso a conteúdos audiovisuais

O Programa de Digitalização de Acervos e Banco de Conteúdos Audiovisuais Brasileiros sistematiza os processos relativos a acervo, licenciamento, digitalização e armazenamento de conteúdos audiovisuais, e construção de interfaces para a Internet, para a propagação e desenvolvimento de modelos de negócio relacionados à convergência digital, integrando políticas de restauração, preservação, catalogação, documentação e difusão.

A expectativa é de que o acesso ao Banco se dê por meio de um portal *online*, nutrido pela Cinemateca Brasileira e outras instituições integrantes do Sistema Brasileiro de Informações Audiovisuais (SiBIA).

O banco abrigará conteúdos audiovisuais digitalizados e seus respectivos dados catalográficos, disponibilizando-os ao público como parte de uma política mais ampla de preservação e acesso do acervo nacional de imagens em movimento.

Também o Centro de Referência Audiovisual – CRA²⁸⁰,

[que] foi concebido para tornar-se o grande repositório de dados sobre conteúdos audiovisuais brasileiros e informações conexas, hoje disponíveis, mas dispersos nos diversos órgãos de cultura da administração federal (Cinemateca Brasileira, Biblioteca Nacional, FUNARTE, etc), em instituições públicas dos estados e municípios do país, bem como nos mais significativos acervos privados.²⁸¹

As atividades do Centro constituem: “Mapear, catalogar e monitorar os conteúdos audiovisuais disponíveis nacionalmente, concentrando o que for possível na

²⁷⁹ <http://www.bcc.org.br/>

²⁸⁰ <http://programadorabrasil.gov.br/cra/>

²⁸¹ FELICE, Fabricio. 2015.

estrutura única de armazenamento”²⁸², bem como procurando por recursos técnicos para um fácil acesso aos acervos e bancos de dados.

A principal meta é “desenvolver conceitos e soluções para um modelo integrado e sustentável de preservação e acesso aos acervos culturais, materializando uma política nacional de digitalização de acervos.”²⁸³

No ano de 2017 também foi lançado um manual da Pinacoteca²⁸⁴.

A Pinacoteca, uma instituição com 110 anos de existência e com a missão institucional de constituir, consolidar e ampliar, estudar, salvaguardar e comunicar um acervo museológico, arquivístico e bibliográfico de artes visuais, produzido por artistas brasileiros ou intrinsecamente relacionado com a cultura brasileira, seus edifícios e memórias, precisa, portanto, estar preocupada com suas informações digitais primitivas. Para tanto, é necessário planejar a aquisição, organização e distribuição de recursos para prevenir uma maior deterioração ou renovar a usabilidade de grupos selecionados de documentos arquivísticos digitais²⁸⁵.

Essa Política de Preservação Digital elaborada pelo Centro de Documentação e Memória da Pinacoteca de São Paulo “procura lançar de forma coordenada as principais bases para que seja implantado, de forma harmoniosa, as novas políticas do processo de preservação de patrimônio documental e artístico digital”²⁸⁶. A iniciativa enxerga os pilares abaixo como essenciais:

- Políticas e normas corporativas de gestão de documentos
- Planos estratégicos de preservação
- Ambiente tecnológico específico (sistema e repositório)²⁸⁷

O projeto contempla todos os tipos de documentos e não apenas os audiovisuais, possui padrões de formatos ou arquivos permitidos e tem objetivos planejados até 2026.

²⁸² FELICE, Fabricio. 2015.

²⁸³ FELICE, Fabricio. 2015.

²⁸⁴ Política de Preservação – PINACOTECA. 2017. p.3.

²⁸⁵ Política de Preservação – PINACOTECA. 2017, p.4.

²⁸⁶ Política de Preservação – PINACOTECA. 2017, p.3.

²⁸⁷ Política de Preservação – PINACOTECA. 2017, p.3.

Existe, ainda, o Dicionário PREMIS,²⁸⁸ que funciona como um esquema e infraestrutura “de metadados voltados para a preservação digital, que teve como maior desafio a antecipação de quais informações são realmente necessárias para suportar um processo específico de preservação”²⁸⁹. Portanto, ele descreve os metadados de preservação, define o conjunto central de metadados (...) que devem ser realizados e identifica/avalia as estratégias para codificação, armazenamento e gerenciamento ou melhorias dos metadados de preservação nos sistemas de arquivamento digital²⁹⁰.

Iniciativas conduzidas pelo IBICT²⁹¹, a CTDE²⁹² e o Arquivo Nacional²⁹³ já visam à “criação de repositórios digitais confiáveis, em conformidade com as normas e padrões vigentes”²⁹⁴, da mesma forma que, por esse motivo, se faz fundamental a reflexão acadêmica e intelectual, bem como

as ações técnicas operacionais direcionadas às questões da salvaguarda e do acesso continuado a representantes nascidos digitais ou digitalizados a partir de itens documentais audiovisuais. Isso porque [considera-se] que difundir o patrimônio audiovisual custodiado por instituições públicas é um passo importante para se alcançar um novo paradigma de relações e de correlações entre as instituições de custódia de documentação audiovisual e seus usuários, produtores, diretores, pesquisadores, estudantes e público interessado, em geral.²⁹⁵

Portanto, na reflexão sobre a produção audiovisual brasileira, podemos perceber que ela vem sendo concebida já em meios digitais “sem necessariamente atender aos requisitos mínimos necessários à sua preservação e, mais ainda, verifica-se uma ausência de políticas públicas voltadas à salvaguarda digital”²⁹⁶.

²⁸⁸ <http://www.loc.gov/standards/premis/>

²⁸⁹ <http://www.brapci.inf.br/index.php/article/view/0000009298/e3a870e31a9e3adbc4c0bacb0ad8e589>

²⁹⁰ <https://www.slideshare.net/redecariniana/preservao-digital-de-arquivos-audiovisuais>

²⁹¹ Instituto Brasileiro de Informação em Ciência e Tecnologia.

²⁹² Câmara Técnica de Documentos Eletrônicos.

²⁹³ <http://arquivonacional.gov.br/br/>

²⁹⁴ SILVA, R. D. et al. 2016, p 1.

²⁹⁵ SILVA, R. D. et al. 2016, p 1.

²⁹⁶ SILVA, R. D. et al. 2016, p 1.

CONSIDERAÇÕES FINAIS

Este trabalho foi produzido com apoio e consentimento da Associação Brasileira de Preservação Audiovisual – ABPA²⁹⁷ que foi fundada em junho de 2008 no 3º Encontro Nacional de Arquivos de Imagem em Movimento, durante 3ª Mostra de Cinema de Ouro Preto – CINE OP. Desde então, cresce o número de interessados e também as pesquisas acadêmicas no ramo. “A ABPA conta atualmente com mais de 200 integrantes: pesquisadores, técnicos, estudantes e profissionais ligados a preservação audiovisual no Brasil”²⁹⁸.

Diante de tamanha fragilidade, esta pesquisa não pretende abraçar todos os filmes produzidos, tampouco todas as realidades de arquivo existentes. Mas, se porventura, o público entender a necessidade real desse tipo de trabalho de constante inovação nesse sentido.

Um artigo²⁹⁹ de especialistas sobre o assunto afirma que uma solução viável para acervos audiovisuais é "a migração de conteúdos para mídias contemporâneas e a atualização de formatos para versões mais recentes"³⁰⁰.

Essa atualização deve procurar atender a dois objetivos: atualização dos suportes e preservação dos conteúdos. De acordo com Dreer, no artigo Preservação de vídeo para milênios, 'estamos vivenciando não apenas uma mudança em como preservar conteúdo em vídeo, mas também uma mudança nas nossas habilidades enquanto preservacionistas e arquivistas'³⁰¹.

Dentre as competências que os arquivistas e funcionários de arquivos e cinematecas precisam adquirir, está o planejamento e execução de ações para a preservação dos digitais. Deve ser avaliada a possibilidade de "manutenção das rotinas de preservação necessárias ao armazenamento, a migração e a atualização de hardware/software, uma vez que a preservação digital é dinâmica e requer ações dirigidas constantes"³⁰².

²⁹⁷ Mais informações no site: <http://abpreservacaoaudiovisual.org/site/>

²⁹⁸ FELICE, Fabricio. 2015.

²⁹⁹ TARTAGLIA, Ana Renata; BUTRUCÉ, Débora. 2015, p.11.

³⁰⁰ TARTAGLIA, Ana Renata; BUTRUCÉ, Débora. 2015, p.11.

³⁰¹ DREER, 2014 apud TARTAGLIA, Ana Renata; BUTRUCÉ, Débora. 2015, p.11.

³⁰² TARTAGLIA, Ana Renata; BUTRUCÉ, Débora. 2015, p.11.

Essa possibilidade da instituição não consiste, apenas, em considerar recursos humanos e financeiros, mas um maior comprometimento com a política de preservação, acima de tudo.

Uma das questões relatadas por Butruce e Tartaglia no mesmo artigo é comum à maioria dos acervos:

assim como as atividades da instituição não cessam, seu acervo audiovisual não para de crescer. E para que se possa continuar a exercer devidamente as atividades de organização, conservação e acesso ao material produzido, é necessária a reflexão constante sobre as estratégias utilizadas para o prolongamento da vida desses materiais, que demandam migrações periódicas e, por conseguinte, metodologia e planejamento integrados e dinâmicos. Ademais, a enorme variedade de suportes e de formatos digitais que não param de surgir atualmente, consequência da permanente transformação tecnológica na área audiovisual, obriga-nos a acompanhar de perto as mudanças e a providenciar medidas seguras que visem a preservação destes conteúdos a longo prazo. E isso vale para todas as outras instituições, grandes ou não.³⁰³

Bem sabemos que o gerenciamento de materiais digitais audiovisuais guardados em servidores já não é mais só um tópico, mas um padrão aplicado, principalmente para a biblioteconomia e ciência da informação. Ambas costumam seguir também o modelo de OAIS. Modelo que serve nesse e em outros casos para aplicação de conceitos, considerando-se metadados e informações do organograma que está no quarto capítulo desse trabalho.

O repositório, como apresentamos, ainda é um lugar considerado de longo prazo no digital, e um dos mais conhecidos é o da Universidade de Hamburgo: este guarda, inclusive, vídeos de pesquisa dos alunos. A UFMG possui iniciativa semelhante, por guardar as produções de alunos do curso de Cinema e afins em um canal no VIMEO.

O digital, em seu amplo contexto, é potente, porém problemático. Precisamos pensar na importância fundamental de preservar o cinema nacional e também as produções menores, bem como as implicações dos padrões e protocolos institucionalizados. Fato é que temos muitos caminhos possíveis, mas em poucos deles temos controle do futuro dos arquivos. Cabe a nós ainda muita discussão e

³⁰³ TARTAGLIA, Ana Renata; BUTRUCÉ, Débora. 2015, p.12.

estudo. Os protocolos aqui colocados não foram descritos em seu passo a passo porque as mudanças de estrutura são muitas e rapidamente o manual ficaria obsoleto.

A área arquivística hoje é a que mais desenvolve padrões para a preservação do digital, e esta pesquisa vem contribuir com informações acerca da necessidade e eficácia da internet nesse processo.

No Brasil, temos também a Cariniana, uma rede brasileira de serviço de preservação digital. Muitos textos e apresentações utilizados na pesquisa foram produzidos por eles, que advertem: toda instituição tem que ter um plano do que fazer com seus arquivos e uma política de preservação, antes de qualquer atividade nesse sentido.

Ao mesmo tempo, vemos vários lugares de guarda utilizando as fitas LTO, explicadas no capítulo 2 dessa dissertação. A tecnologia já está em sua versão de número 7 já caminhando para a 8. Quem começou a preservar assim já não está mais conseguindo acompanhar, afinal, são gastos obrigatórios e uma migração forçada após 2 versões.

Esta dissertação não tem como objetivo trabalhar a questão dos direitos e da autoria no contexto da internet, que é delicada, e careceria de acordos e deliberações por parte das instituições e também dos produtores do conteúdo.

O ponto de vista — polêmico — de “antes guardar o arquivo tal como ele é e foi concebido do que propor parâmetros quase inatingíveis” também não foi muito explorado, mas deve ser lembrado. Como ficariam, por exemplo, o caso dos conteúdos amadores, dos cinemas marginais? Com a democratização do audiovisual que já foi explicitada no capítulo 3, temos muita coisa sendo produzida e é realmente difícil acompanhar e preservar tudo. Em seu mestrado³⁰⁴, Inês Aisengart fez uma proposta para a preservação dos materiais audiovisuais amadores gravados

³⁰⁴ MENEZES, Inês. Aisengart. Preserving amateur digital records of public demonstrations in 2013/2014 in Rio de Janeiro.

durante as manifestações de 2013, produzidas por “Mídia Ninja³⁰⁵” e mesmo cinegrafistas amadores que se aventuraram com suas câmeras. São materiais de enorme relevância no contexto da comunicação brasileira e se faz extremamente necessário que nossa geração compreenda o valor desses arquivos e a importância que eles terão, principalmente no futuro. Para tanto, eles também precisam ser lembrados quando se discute preservação audiovisual como um todo.

Vivemos num país que passa por inúmeras transformações políticas e as provas em áudio e vídeo são grandes tesouros a dizer de uma época. De delações a manifestações nas ruas, todo material produzido é de extrema relevância para contar essa história daqui alguns anos.

A dissertação também não entrou em detalhes sobre a criação de nenhuma plataforma nova nesse sentido. Mas penso que seria incrível se uma plataforma, como o próprio Netflix, fosse criada ou aproveitada no Brasil para arquivamento de seu conteúdo (principalmente nacional) de forma que garantisse o acesso do usuário — principal foco desta pesquisa. Durante o governo Dilma, uma proposta semelhante foi pensada, com investimento inicial de 10 milhões³⁰⁶. Grandes foram as críticas por parte da população por pensarem que num país como o nosso, onde falta o básico (como desenvolvimento, educação e saúde), esse dinheiro seria mais bem gasto em outro lugar. De fato, uma pessoa “com fome” e sem acesso à própria internet acharia inútil ter “Limite”, de Mário Peixoto disponível para assistir. Mas, se pensarmos dessa forma, nada nunca será feito.

Dessa forma, também é importante ressaltar que não temos poder de decisão sobre o que deve ou não ser preservado. Podemos não gostar do filme novo da Globo Filmes, mas também ele deve ser preservado: antes do juízo de valor deve vir a consciência de que essa grande empresa domina o mercado produtivo e diz sobre algo cultural da contemporaneidade que vivemos. Da mesma forma, o filme produzido sem dinheiro, independente, por estudantes de cinema, amantes,

³⁰⁵ Canal online não oficial de comunicação que busca transparência em relação às mídias oficiais (canais de TV, jornais, rádios e etc).

³⁰⁶ <https://fernandorodrigues.blogosfera.uol.com.br/2016/03/11/governo-dilma-gasta-r-10-milhoes-para-criar-netflix-brasileiro-estatal/>

amadores que se arriscam nas ferramentas e na linguagem, também exalam importância cultural e merecem ser igualmente preservados.

Importante acentuar que precisamos, atualmente, melhorar os padrões. Um exemplo disso é preservar não apenas os filmes, mas também os arquivos correlatos aos filmes, tais como: roteiros, cartazes, cenas brutas não utilizadas, dentre outros.

Portanto, a principal função de estarmos debatendo esse assunto é manter os documentos audiovisuais acessíveis e disponíveis, afinal a digitalização traz questões e novos problemas. E toda questão ligada à materialidade desses arquivos pode ser compreendida em relação ao fato de que existe materialidade no documento digital, ele está fisicamente em algum lugar, mesmo na nuvem, pois esta demanda um servidor. Dessa forma, concluímos que a Internet e redes aumentam as possibilidades, e que os arquivos deviam considerar que agora temos mais recursos e ferramentas.

Este trabalho também entende que o valor histórico e documental de um vídeo ou filme é muito maior que a técnica utilizada, e que, às vezes, é necessário tirar o foco das especificações técnicas para levantar alguns questionamentos e definir caminhos mais claros dentro do digital.

Da mesma forma, não sejamos omissos e não trabalhem com uma presunção de autenticidade — as tecnologias e procedimentos podem garantir identidade e integridade do arquivo, mas não se deve descuidar da noção de autenticidade que lhe é intrínseca.

REFERÊNCIAS

- ASSMAN, Aleida. Sobre as metáforas da recordação. In: **Espaços da recordação; formas e transformações da memória cultural**. Trad. Paulo Soethe (coord). Campinas: Ed. da Unicamp, 2011, p. 161-170; p. 221-233.
- BEIGUELMAN, Giselle. **Reinventar a memória é preciso**. In: *Futuros possíveis: Arte, Museus e Arquivos Digitais*. São Paulo: Peirópolis, 2014.
- BENJAMIN, Walter. **Desempacotando minha biblioteca**. In: *Obras escolhidas II*. Rua de mão única. 5ª. ed. São Paulo: Brasiliense, 1985.
- BENJAMIN, Walter. **Passagens**. São Paulo: Iluminuras, 1998.
- BENJAMIN, Walter. **O anjo da história**. Belo Horizonte: Autêntica Editora, 2012. P. 7-20
- BENJAMIM, Walter. **Origem do drama barroco alemão**. Trad. Sérgio Paulo Rouanet. São Paulo: Brasiliense, 1984. “Ruína” – p. 199 a 204.
- BENJAMIM, Walter. Sobre o conceito de história. In: **Magia, técnica, arte e política; ensaios sobre literatura e história da cultura**. Obras escolhidas, Volume 1. São Paulo: Brasiliense, 1987, p. 222-232.
- BESSER, Howard. **Longevidade Digital**. Revista Acervo. Rio de Janeiro, 2010. p.58
- BESSER, Howard. **Digital Preservation of Moving Image Material?**. 2001. UCLA School of Education & Information Studies. Disponível em: <<http://besser.tsoa.nyu.edu/howard/Papers/amia-longevity.html>>. Acesso em: 01 mar 2018.
- BRADLEY, Rachael. **Digital Authenticity and Integrity: Digital Cultural Heritage Documents as Research Resources**. Libraries and the Academy. Disponível em: <http://muse.jhu.edu/journals/portal_libraries_and_the_academy/v005/5.2bradley.pdf> Acesso em: 01 mar 2018.
- BEZERRA, Laura. In: **Catálogo de Reflexões sobre a Preservação Audiovisual: 10 anos de Cine OP**. 2016
- BURGESS, Jean e GREEN Joshua. Youtube e a revolução digital. 2009 Disponível em: <<https://goo.gl/F2xG4q>> Acesso em: 08 mar 2018.
- CICCARINI, Rafael. Centro Universitário UNA - aula. Anotações da autora, 2012.
- CASTELLO BRANCO, Lucia. **A traição de Penélope**. São Paulo: Annablume Editora, 1994.
- CINEMATECA BRASILEIRA. **Página inicial**. Disponível em: <<http://www.cinemateca.gov.br>>. Acesso em: 09 jan. 2017.

CINEMATECA DO MAM. **Página principal**. Disponível em: <http://mamrio.org.br/wp/museu_cinemateca/apresentacao>. Acesso em: 09 jan. 2017.

CINEMATECA FRANCESA. **Página principal**. Disponível em: <<http://www.cinematheque.fr/>>. Acesso em: 05 jan. 2017.

CUEVAS ÁLVAREZ, Efrén. **La memoria del cine o el cine que recuerda**, en Josep M. Català (ed.), *El cine de pensamiento: Formas de la imaginación tecnostética*, Colección Aldea Global, UAB / UPF / UJI / UV, 2014, pp. 189-203.

DANTO, Arthur C. **Looking at the future looking at the present as past**. In: Miguel Angel Corzo (ed.), *Mortality Immortality?: The Legacy of 20th-Century Art*, Getty Conservation Institute (Los Angeles, CA), 1999.

DERRIDA, Jacques. **Mal de arquivo: uma impressão freudiana**. Trad. Cláudia de Moraes Rego. Rio de Janeiro: Relume Dumará, 2001.

DIDI-HUBERMAN, Georges. **Quando as imagens tocam o real**. In: Pós, Vol. 2, no. 4, p.204-219, nov. 2012.

DIOGO, Ligia Azevedo. **Vídeos de família: entre os baús do passado e as telas do presente**. Dissertação de mestrado. Niterói: UFF, Universidade Federal Fluminense, 2010.

DPC ONLINE. **Manuais de preservação**. Disponível em: <<https://dpconline.org/handbook>>. Acesso em: 01 mar. 2018.

EDMONDSON, Ray. **Filosofia e princípios da arquivística audiovisual**. UNISIST: United Nations Educational, Scientific and Cultural Organization, tradução de 1998. Disponível em: https://edisciplinas.usp.br/pluginfile.php/192381/mod_resource/content/3/EDMONSONfilosofiadearquivosAudiovisuais.pdf. Acesso em: 07 ago 2017.

EDMONDSON, Ray. **Filosofia e princípios da arquivística audiovisual**. Tradução de Carlos Roberto de Souza. Rio de Janeiro: Associação Brasileira de Preservação Audiovisual; Cinemateca do Museu de Arte Moderna do Rio de Janeiro, 2013. 224 p.

EDMONDSON, Ray. **Filosofia e princípios da arquivística audiovisual**. UNISIST: United Nations Educational, Scientific and Cultural Organization, Trad. de Carlos Roberto Rodrigues de Souza. – Brasília : UNESCO, 2017. Disponível em: <http://unesdoc.unesco.org/images/0025/002592/259258POR.pdf> Acesso em: 01 mar 2017.

ERNST, Wolfgang. **Digital memory and the archive**. Minneapolis: University of Minnesota Press, 2013.
Archageography: method and machine versus the history and narrative of media.

FELICE, Fabrício. **Ações para a preservação digital de acervos audiovisuais no Brasil**. Apresentação de slides. Fórum 2015. Disponível em: <http://forum2015.rnp.br/sites/forum2015/files/forum_rnp2015_apresentacao_fabricio_felice.pdf>. Acesso em: 01 mar. 2018.

FERREIRA, Miguel. **Introdução à preservação digital**, 2006.

FONTANA, F. F. et al. **Archivematica como ferramenta para acesso e preservação digital à longo prazo**. ÁGORA, Florianópolis, v. 24, n. 48, jul. 2014. Disponível em: <http://www.brapci.inf.br/_repositorio/2014/12/pdf_4de91da221_0028448.pdf>. Acesso em: 01 mar. 2018.

FOSTER, Lila Silva. **Filmes domésticos**: uma abordagem a partir do acervo da Cinemateca Brasileira. Dissertação de mestrado. São Carlos: UFSCar, Universidade Federal de São Carlos, 2010.

FRANÇA, Andréa. In: **Catálogo de Reflexões sobre a Preservação Audiovisual: 10 anos de Cine OP**. 2016

GOMES, Paulo Emílio Salles Gomes. **Cinema**: trajetória no subdesenvolvimento. Rio de Janeiro: Paz e Terra, 1980.

GUASCH, Anna Maria. **Arte y archivo, 1920-2010**, Genealogías, tipologías y discontinuidades. Madrid: Akal, 2010. Capítulo 3, p. 46-64; 97-107.

HERMANN, Leticia. A convergência midiática e as mudanças comportamentais no consumo do mercado de nicho: Netflix e a “desmaterialização” dos produtos. **Revista Interamericana de Comunicação Midiática**, [S.L], v. 11, n. 22, p. 222-245, jul. 2012. Disponível em: <<https://periodicos.ufsm.br/animus/article/viewFile/7080/pdf>>. Acesso em: 01 mar. 2017.

HOFF, Rafael Sbeghen. **Viral: ensaio sobre correlações entre dispositivos móveis e experiência estética audiovisual**. 2015. Disponível em: <<http://portalintercom.org.br/anais/nacional2015/resumos/R10-2433-1.pdf>> Acesso em: 01 mar 2018

INFOGRÁFICO. **Dados sobre o acesso de internet no Brasil**. Disponível em: <http://blog.pmweb.com.br/wp-content/uploads/infografico_internet_no_brasil.png>. Acesso em: 29 abr. 2017.

JAKOBSEN, Kjetil. **Anarchival society**. In: ROSSAAK, Eivind. **The archive in motion**: new conceptions of the archive in contemporary thought and new media practice. Oslo: Novus Press, 2010.

KITCHIN, Rob. **Thinking critically about and researching algorithms**. 2014. Disponível em http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2515786. Acesso em 18 mar 2016.

LA CARRETTA, Marcelo L. da Cunha; (or: Nazario, Luiz). **Cinema: memória audiovisual do mundo**. 2005. Dissertação (Mestrado) - Universidade Federal de Minas Gerais. Disponível em: <<http://www.bibliotecadigital.ufmg.br/dspace/handle/1843/VPQZ-73BJW9>>. Acesso em: 15/06/2015.

LA CARRETTA, Marcelo. **Prosumidores nostálgicos: reflexões sobre a obsolescência das mídias audiovisuais**. 2012. 232 f. Tese (Doutorado em Artes) – Escola de Belas Artes, Universidade Federal de Minas Gerais, Belo Horizonte, 2012.

LA CRUZ, Sonia Estela Montañó. **Plataformas de vídeo: apontamentos para uma ecologia do audiovisual da web na contemporaneidade**. São Leopoldo: Unissinos, 2012. Tese de Doutorado apresentada ao Programa de Pós-Graduação em Ciências da Comunicação da Universidade do Vale do Rio dos Sinos. Disponível em: <<http://www.repositorio.jesuita.org.br/bitstream/handle/UNISINOS/3478/Sonia+Monta%F1o.pdf;jsessionid=F5D627EBCEAFBB265CE1577A2FF2B01F?sequence=1>> Acesso em 01 mar 2018.

MARCONI, Marina de Andrada; LAKATOS, Eva Maria. **Fundamentos da Metodologia Científica**. São Paulo, Atlas, 2003, 311 p.

MASCARELLO, Fernando. **História do Cinema Mundial**. Campinas: Papyrus, 2006.

MENEZES, Inês. **O digital não vai esperar**. In: **Catálogo de Reflexões sobre a Preservação Audiovisual: 10 anos de Cine OP**. 2016

MENEZES, Inês. **Aisengart. Preserving amateur digital records of public demonstrations in 2013/2014 in Rio de Janeiro**. UvA Scripties, [S.L], mar. 2015. Disponível em: <<http://www.scriptiesonline.uba.uva.nl/531351>>. Acesso em: 01 mar. 2018.

MENEZES, Roniere. **Mário de Andrade: arquivo, tradução e traição da memória**. In: SAID, Roberto e Sá, Luiz Fernando (Orgs.). *Jacques Derrida: entreatos de leitura e literatura*. Cotia, SP: Ateliê Editorial, 2014.

MENEZES, Ulpiano T. Bezerra de. **Memória e cultura material: documentos pessoais no espaço público**. Revista Estudos Históricos, Rio de Janeiro, v.11, n.21, 1998, p.89-103.

MUSSA, Ivan; et al. BARBOSA, Rafael; GARCIA, Yuri. **Resgate midiático na Internet - Democratização e produção no mercado de nicho online**. XI INTI International Conference La Plata, [S.L], out. 2012. Disponível em: <http://www.memoria.fahce.unlp.edu.ar/trab_eventos/ev.2674/ev.2674.pdf>. Acesso em: 01 mar. 2018.

ACADEMY OF MOTION PICTURE ARTS AND SCIENCES. **O dilema digital 1: questões estratégicas na guarda e no acesso a materiais cinematográficos digitais**. São Paulo: Cinemateca Brasileira, 2009.

ACADEMY OF MOTION PICTURE ARTS AND SCIENCES. **O dilema digital 2:** perspectivas de cineastas independentes, documentaristas e arquivos audiovisuais sem fins lucrativos / Traduzido por Millard Schisler; Osvaldo Emery e Patricia de Filippe. São Paulo: Instituto Butantan, 2015.

OBSERVATÓRIO DA DIVERSIDADE. **Mudança do MIS.** Disponível em: <<http://observatoriodadiversidade.org.br/site/prefeitura-de-belo-horizonte-determina-mudanca-de-endereco-do-museu-da-imagem-e-do-som/>>. Acesso em: 13 fev. 2018.

PALACIOS, Marcos Silva. **A memória na era Digital.** Palestra na abertura do 10º ALCAR – Encontro Nacional de História da Mídia, em 03 jun 2015, Porto Alegre, Transcrição minha.

PIETROBRUNO, Sheenagh. **YouTube and the social archiving of intangible heritage.** 2013 Disponível em: <http://citereerx.ist.psu.edu/viewdoc/download?doi=10.1.1.902.2956&rep=rep1&type=pdf> Acesso em: 01 mar 2018

PINACOTECA. **Política de preservação.** Disponível em: <<http://pinacoteca.org.br/wp-content/uploads/2016/11/politica-de-preservacao-pinacoteca2017-versao-externa.pdf>>. Acesso em: 01 mar. 2018.

QUENTAL, José. 2016, p. 36. In: **Catálogo de Reflexões sobre a Preservação Audiovisual: 10 anos de Cine OP.** 2016

REIS, Isabella Rodrigues. **Netflix nas Redes Sociais: Uma Estratégia Pontual.** Portal Intercom, Uberlândia, jun. 2017. Disponível em: <<http://portalintercom.org.br/anais/sudeste2017/resumos/R58-0724-1.pdf>>. Acesso em: 01 mar. 2018.

RESOLUÇÃO DO MEC. **Disciplina de Preservação.** Disponível em: <http://portal.mec.gov.br/cne/arquivos/pdf/rces10_06.pdf>. Acesso em: 20 nov. 2017.

RESOLUÇÃO DO MEC. **Preservação como disciplina obrigatória.** Disponível em: <<http://portal.mec.gov.br/busca-geral/323-secretarias-112877938/orgaos-vinculados-82187207/12709-resolucoes-ces-2006>>. Acesso em: 20 nov. 2017.

REVISTA ENCONTRO. **Mudança do MIS.** Disponível em: <<https://www.revistaencontro.com.br/canal/atualidades/2017/05/mudanca-de-endereco-do-museu-da-imagem-e-do-som-e-criticada-por-veread.html>>. Acesso em: 13 fev. 2018.

RICOEUR, Paul. **A memória, a história, o esquecimento.** Tradução Alain François et. al. Campinas, SP: Editora da Unicamp, 2007.

RICOEUR, Paul. **Tempo e narrativa**, vol. 3. Trad. Roberto Leal Ferreira. Campinas: Papyrus, 1997, p. 196-216.

RODRIGUES, Eloy. 2006, p.11-12. In: [Prefácio] FERREIRA, Miguel. **Introdução à preservação digital**, 2006.

SANTAELLA, Lucia. **Cultas e arte do pós-humano: da cultura das mídias à cibercultura**. São Paulo: Paulus, 2003.

SIGNIFICADOS. **Significado de binário**. Disponível em: <<https://www.significados.com.br/binario/>>. Acesso em: 20 nov. 2017.

SILVA, Armando. **Álbum de família: a imagem de nós mesmos**. Trad. Sandra Martha Dolinsky. São Paulo Editora SENAC São Paulo; Edições SESC São Paulo, 2008.

SILVA, R. R.G. **Desafios e Alternativas Digitais para a Salvaguarda e Difusão do Patrimônio Público Documental Arquivístico Audiovisual**. Slides apresentados no I SINPRED – Seminário Internacional de Preservação Digital; III Encontro Nacional da Rede cariniana. Instituto Brasileiro de Informação em Ciência e Tecnologia – IBICT, Brasília, 07 a 09 de maio de 2014, 2014a. Disponível em: http://cariniana.ibict.br/index.php/publicacoes/item/download/121_c92f08b70ebb4910fcca97cf492cfeec . Acesso em: 28 abr 2017.

SILVA, R. R.G. **Digitalização de acervos fotográficos públicos e seus reflexos institucionais e sociais: tecnologia e consciência no universo digital**. Tese. Ciência da Informação. Orientadora: Lena Vania Ribeiro Pinheiro. Rio de Janeiro, PPGCI; UFRJ/ECO; IBICT, 2002. Disponível em: . Acesso: em 01 mai 2018.

SILVA, R. R.G. **Patrimônio documental audiovisual em arquivos de instituições públicas soteropolitanas**: prosseguindo na pesquisa e expandindo-a às cidades de São Paulo, Rio de Janeiro e Belo Horizonte. Anais do IX ENECULT – Encontros de Estudos Multidisciplinares em Cultura, 11 a 13 de setembro de 2013, 2013e, Salvador (BA). Disponível em: rel_download.asp?nome=48263.pdf>. Acesso em: 01 mai 2018.

SILVA, R. R.G. **Patrimônio documental audiovisual em instituições públicas: Desafios e alternativas**. Anais do 9º CineOP – Mostra e Cinema e Outro Preto / 9º Encontro de Arquivos Nacionais e Acervos Audiovisuais Brasileiros, 28 de maio a 2 de junho de 2014, 2014b, Ouro Preto (MG).

SILVA, R. D. et al. **A iniciativa legatum e a preservação digital de arquivos audiovisuais públicos**. Periódicos Unicamp, [S.L], v. 14, n. 3, set./dez. 2016. Disponível em: <<https://periodicos.sbu.unicamp.br/ojs/index.php/rdbci/article/view/8646279/pdf>>. Acesso em: 01 mar. 2018.

SILVA, Sabrina Tenório Luna da. **Subjetividade, arquivos familiares e found footage no documentário brasileiro contemporâneo**. 2012. Artigo (Doutorado) – Universidade Federal de Pernambuco. Disponível em: <www.doc.ubi.pt/13/dossier_sabrina_silva.pdf>. Acesso em: 15 mai 2015.

SONY. **Tecnologia oda**. Disponível em: <<https://www.sonypro.com.br/pro/article/broadcast-products-sony-oda>>. Acesso em: 13 fev. 2018.

STECZ, Solange Straube In: **Catálogo de Reflexões sobre a Preservação Audiovisual**: 10 anos de Cine OP. 2016

TARTAGLIA, Ana Renata; BUTRUCE, Débora. **Entre letras e imagens**: o acervo audiovisual no Arquivo da Academia Brasileira de Letras. Revista AGCRJ, Rio de Janeiro, n. 9, mar. 2015. Disponível em: <http://wpro.rio.rj.gov.br/revistaagcrj/wp-content/uploads/2016/11/e09_a19.pdf>. Acesso em: 01 mar. 2018.

VELOSO, Geraldo. In: **Catálogo de Reflexões sobre a Preservação Audiovisual**: 10 anos de Cine OP. 2016

AUDIOVISUAL

CINEMA Paradiso. Direção e Roteiro: Giuseppe Tornatore. Produção: Franco Cristaldi. Intérpretes: Philippe Noiret; Jacques Perrin; Antonella Attiu; Enzo Cannavale; Isa Daniello, Leo Gullota; Marco Leonardi; Puppella Maggio, Leopoldo Trieste; Salvatore Cascio e outros. Música: Ennio Morricone. Roma: Cristaldifilm, 1990. DVD (123 min), son., color.

Documentário Salvadores de imagens (Keepers of the frame, 1999), da WinStar Productions.

APÊNDICE A – Perguntas para composição do questionário

CONSUMO AUDIOVISUAL

Este questionário foi produzido para uma pesquisa acadêmica, precisamos de sua ajuda e sinceridade. É rápido! :)

* Required

Qual a sua faixa de idade? *

- Menor de 18 anos
- 18 a 24 anos
- 25 a 39 anos
- Mais de 40 anos

Qual o seu nível de escolaridade? *

- Ensino Fundamental - incompleto
- Ensino Fundamental - completo
- Ensino Médio/Técnico - incompleto
- Ensino Médio/Técnico - completo
- Ensino Superior - incompleto
- Ensino Superior - em andamento
- Ensino Superior - completo
- Pós-graduação / Mestrado / Doutorado - incompleto
- Pós-graduação / Mestrado / Doutorado - completo

Você trabalha? *

- Sim, tempo integral
- Sim, meio horário
- Sim, faço estágio
- Sim, como freelancer
- Sim, como autônomo
- Não

Qual a sua profissão? *

Your answer

Você tem o hábito de consumir conteúdo audiovisual?
(filmes, séries e afins...) *

- Sim, muito
- Sim, mas pouco
- Não

Marque as alternativas que melhor representam seus
hábitos (se houverem): *

- Ir ao cinema
- Ver filmes/séries na TV
- Ver filmes/séries em outros dispositivos (celular, tablet, computador...)
- Participar de Mostras e Festivais de Cinema

Que tipo de conteúdo você prefere? *

- Filmes nacionais
- Filmes estrangeiros
- Filmes de animação
- Filmes live-action
- Séries

Você baixa filmes na internet? (torrent, etc.)

- Sim, sempre
- Sim, às vezes
- Não

Você assina TV paga? (sky, net, claro, etc.) *

- Sim
- Não

Você assina Netflix? *

- Sim
- Não

Você já assistiu filmes (longa duração) pelo YouTube? *

- Sim
- Não

Já contratou filmes por demanda? (Net NOW, Google Play, Vivo Play, Sky Online, Mubi, etc...) *

- Sim
- Não

Já ficou sem encontrar algum título que desejava na internet? *

- Sim, muitas vezes
- Sim, mas poucas vezes
- Não

Você compra DVD ou Blu-ray dos filmes/séries que gosta? *

- Sim, sempre
- Sim, às vezes
- Não

Você frequenta eventos da área audiovisual? (Mostras, Festivais e afins). *

- Sim, sempre
- Sim, às vezes
- Não

Você considera importante preservar e guardar filmes? *

- Sim
- Não
- Não tenho opinião formada

Você produz conteúdo audiovisual? (Fotos, vídeos, áudios).

*

- Sim, com meu celular
- Sim, com câmera amadora
- Sim, com equipamento profissional
- Trabalho com isso
- Não

Você tem o hábito de postar/compartilhar esses conteúdos? *

- Sim, todos
- Sim, mas não todos
- Não

Quais plataformas você utiliza para compartilhamento desses dados? *

- Facebook
- Instagram
- Twitter
- YouTube
- Vimeo
- Google Plus
- Snapchat
- Whatsapp / Messenger / Telegram
- Não compartilho
- Other: _____

Onde você salva seus arquivos? (Fotos, vídeos, áudios, documentos). *

- No computador
- No pendrive / HD externo
- Nas plataformas de compartilhamento (Facebook, Youtube, etc)
- No cartão de memória ou memória interna do celular/tablet
- Faço backup pelo e-mail
- Faço backup em nuvem

Você tem o hábito de fazer backup em nuvem? (Google Drive, Dropbox, Mega, iCloud, One Drive, etc.) *

- Sim, faço sempre
- Sim, faço de vez em quando
- Sim, já fiz uma vez ou outra
- Não

Você confia na segurança dos arquivos na internet? *

- Sim, totalmente
- Sim, com ressalvas
- Tenho receio, mas coloco mesmo assim
- Não

Você consegue ver diferença de qualidade nas imagens abaixo? *



IMAGEM 1

IMAGEM 2

- A imagem 1 tem maior qualidade
- A imagem 2 tem maior qualidade
- Ambas são iguais em qualidade

Você consegue ver diferença de qualidade nas imagens abaixo? *



IMAGEM 1



IMAGEM 2

- A imagem 1 tem maior qualidade
- A imagem 2 tem maior qualidade
- Ambas são iguais em qualidade

Você consegue ver diferença de qualidade nas imagens abaixo?

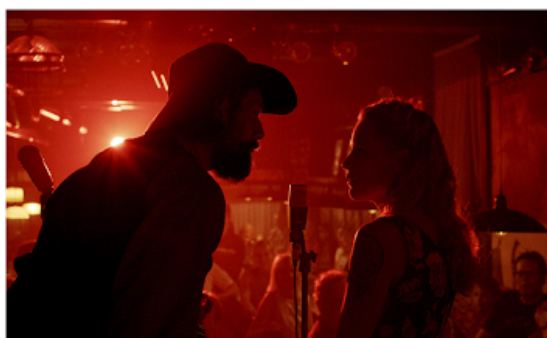


IMAGEM 1

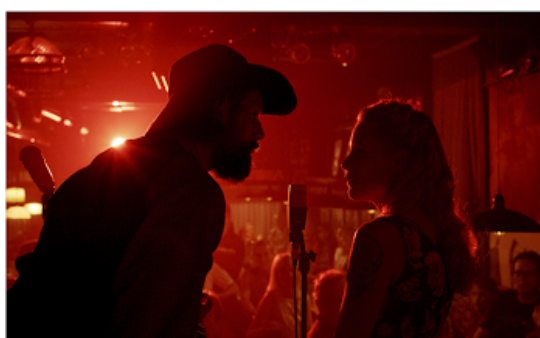


IMAGEM 2

- A imagem 1 tem maior qualidade
- A imagem 2 tem maior qualidade
- Ambas são iguais em qualidade

SUBMIT

ANEXO A – Cartilha DPC online de preservação

Digital Preservation **Handbook**

Getting Started



Illustrations by Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Gold sponsor



Silver sponsors



Bronze sponsors



Reusing this information

You may re-use this material in English (not including logos) with required acknowledgements free of charge in any format or medium. See [How to use the Handbook](#) for full details of licences and acknowledgements for re-use.

For permission for translation into other languages email: handbook@dpconline.org

Please use this form of citation for the Handbook: Digital Preservation Handbook, 2nd Edition, <http://handbook.dpconline.org/>, Digital Preservation Coalition © 2015.

Contents

Getting Started 4

 Introduction..... 4

 Get to Know Your Organisation and Your Data..... 4

 Where next? 7

 Resources 8

 Case studies..... 10

 References..... 11

Getting Started



Illustration by Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Introduction

This section is for you if you have yet to start digital preservation or have just begun to do so. It provides a rapid introduction to a number of approaches that will support you in getting started, moving towards using other sections of the Handbook, and in building confidence and skills.

The section developed out of the "Getting Started in Digital Preservation" workshops run by the Digital Preservation Coalition. It supports 'learning by doing' and assumes a minimal level of prior knowledge. As you progress you will benefit from dipping into the resources and case studies, other topics and sections in the Handbook, and the [Glossary](#) for any unfamiliar terms.

Digital preservation can seem like a daunting prospect. It can help to map out the skills and resources you have and the materials you want to preserve. That way you start with what you know rather than what you don't. The first step in digital preservation is almost always to undertake a rapid assessment. This will have two or three components:

- knowing the practical capacity of your organisation;
- understanding the organisation's goals and missions; and
- knowing a little about the digital materials in question.

Get to Know Your Organisation and Your Data

Creating a Digital Asset Register

As part of a rapid assessment: it is vital to understand the nature and extent of your digital collections. A digital asset register will be incredibly useful for assessing the extent and significance of the collection, identifying priorities and planning digital preservation actions. A high-level assessment of the collection will help with more detailed mapping later: a comprehensive and detailed audit could be time consuming. So the advice in early stages is to keep the asset register simple. Ask the following questions:

- What is the subject of the collection?

- Where does it come from and what is its function?
- Where is it stored and what kinds of media are used?
- Why is it being retained?
- Who is responsible for it; who are the users; who are the subjects of the data?
- How is the data accessed?
- How is the data likely to change and grow in the near future?

Assessing Your Organisation's Readiness

Organisational maturity is another factor to consider. The National Digital Stewardship Alliance ([NDSA, 2013](#)) in the USA has recommended a simple 4 level model to help organisations understand and improve their technical capacity in digital preservation. The four levels are:

- Level 1 - protect your data
- Level 2 - know your data
- Level 3 - monitor your data
- Level 4 - repair your data

These 'Levels of Preservation' are intended to be progressive, and are used to measure maturity against four components: storage, file fixity, information security, metadata, and file formats. An organisation's capacity to undertake digital preservation is indicated by its maturity level across these five components. More comprehensive maturity models are available, such as the Digital Preservation Capability Maturity Model ([Dollar and Ashley, 2014](#)), if a more well-rounded exploration of organisational maturity is required.

First Steps to Securing Your Data

This section provides an overview of initial actions to secure your data once you have assessed your organisation's readiness and compiled basic information about your data. The following steps are essential to ensuring a minimum level of preservation when a new collection of digital material is received. This is typically referred to as bit preservation. Quite literally, preserving the streams of binary digits, or bits, that make up your digital files (without preserving the means to decode the bits into meaningful information).

Prompt check in on receipt

When a new collection of digital material is received from a supplier it is essential to ensure that what has been received is what is expected. Depending on the source of the material, it may be possible to request new copies of missing or poor quality files. These checks are made and any replacement requests submitted, the greater the likelihood of successful resolution.

Key tasks include:

- Scan for viruses and malware to make sure there are no unwanted surprises in the collection. Perhaps keep the collection 'in quarantine' until you have checked it.
- Check all expected files are present. If the material is accompanied by a manifest, check the files against it

- Open a random selection of files to verify their integrity and/or expected quality levels
- Promptly request replacements for any damaged or missing files, where possible

Create a Verifiable File List

In order to check over time that your digital files are being preserved, it is first necessary to record exactly what files are in your possession. It is, therefore, important to create a verifiable list of files in each collection. These lists should likely contain information such as file names, locations and sizes, format types and checksums. A checksum is a short alphanumeric string that represents the contents of a file which acts as a 'digital fingerprint' allowing comparison over time. Once the list has been created, it is a simple process to verify that all files are present and undamaged, at any point in the future (see fixity checking below).

Various software [tools](#) can be used to automatically generate this data; these are commonly referred to as characterisation tools. For example, you may wish to use The National Archives [PRONOM](#) (a register of file formats and their behaviours) and [DROID](#) (a tool that uses PRONOM to analyse the files on a system). Having a list of the file formats, versions and quantities in your collections will help you make a case to senior management for the support and resources that will be needed to do the job properly and sustainably. This information can also be used to update and enrich your digital asset register. The range of formats in use should be consolidated to minimize duplication and eliminate problem formats. This process is known as normalization.

Key tasks:

- Generate a verifiable file list
- Update digital asset register

Stabilise your files: make copies

No matter how good your digital storage, your digital material will always be at risk of damage, decay or accidental deletion. Making more than one copy of your digital materials and utilising more than one type of storage solution mitigates a variety of digital preservation risks.

Key tasks include:

- Keep (at least) one copy easily accessible on non-removable disk. You will need to regularly revisit your material to ensure its fixity, so keeping it accessible will make this easier
- Make (at least) one additional copy, if necessary on a less accessible, but cheaper storage medium such as tape
- Keep one copy in a different geographical location to the others to mitigate against disaster

Revisit and inspect: Fixity checking

By revisiting your digital materials on a regular basis (e.g. every 6 months) you can ensure that no damage or accidental loss has occurred. If it has, you can recover problematic files from the copies or backups you have made previously. Future fixity checks will generate new digital fingerprints (or checksums) for the files in your collections. If they do not match the ones originally created, bit loss or damage has occurred.

Key tasks include:

- Revisit your collection on a frequent basis, recalculate the checksums, identify files that have become damaged

- Retrieve copies of damaged files and repair as necessary
- Perform test recoveries of data backed up by third party services, to ensure backups are being performed as agreed

Document your processes

From the outset of creating a digital collection, it is important to document as much as possible about a collection's assets, the tools and workflows. This documentation is an important component of [technical and descriptive Metadata](#). It is necessary to retain this information for the purposes of longevity. As, with any project, staff retention can be an issue. If staff leave they often take essential knowledge and skillsets.

Where next?

Having taken the first steps in digital preservation, where do you go next? This will obviously depend on your own requirements and priorities, but this table provides a number of suggestions and other sections of the Handbook will help you move forward with them:

Next steps
Develop advocacy and outreach, an understanding of risk , the business case, costs, benefits and impact
Establish an organisational preservation strategy and policies . As well as ensuring a consistent approach to preservation it can be a useful tool to achieve buy in across an organisation and in particular with senior management
Establish a digital repository. Technical solutions and tools either on local IT infrastructure or offered as a cloud service will help you understand, manage and preserve your digital material for the long term
Establish your long-term storage , preservation planning and action
Revisit and expand your collection audits: <ul style="list-style-type: none"> • Characterise priority collections in more detail • Periodically update collection audits as required
Establish a digital preservation working party. Effective digital preservation often requires buy in across many departments within an organisation. A representative working party can be vital in making coordinated steps forward
Build the necessary staff training and development and skill sets
Establish a professional network and collaborations . Join a digital preservation membership organisation such as the Digital Preservation Coalition

Keep up to date with new developments:

- Email lists for digital preservation include the [digital preservation](#) announcement list on JiscMail, and the USA-focussed [digipres](#) list
- A weekly [DP News](#) blog selects recent tweets and news links on digital preservation
- Journals with a digital preservation focus include: [International Journal of Digital Curation](#), and [D-Lib](#)
- Events with a digital preservation focus include [iPRES](#), and [PASIG](#)
- The [Digital Preservation Coalition organises briefing days on particular digital preservation topics](#)

Resources



A Preservation Primer

<http://knconsultants.org/a-preservation-primer/>

This clear practical short primer on preservation for beginners was written by staff at Portico. It summarizes the issues and outlines various short and long-term preservation options that an organization might take to begin planning for long-term digital preservation of its content, beginning with near-term protection and concluding with full preservation and long-term protection. (83 pages).

Don't Panic: The Archivist's Guide to Digital Preservation

http://www.wyjs.org.uk/documents/archives/dont_panic_digital_preservation_first_steps_guide.pdf

A practical and concise guide produced in 2011 by Stefanie Davidson at the West Yorkshire Archive Service. It is intended to act as a signpost to assist in taking the first steps in understanding some of the issues involved rather than a comprehensive guide and an introduction to the topic to help you find your feet. (8 pages).

Putting Parsimonious Preservation into Practice

<http://www.nationalarchives.gov.uk/documents/information-management/parsimonious-preservation-in-practice.pdf>

The principle of Parsimonious Preservation was originally developed in 2009 at The National Archives in the UK as an approach for small or medium sized institutions to permit them to begin work on digital preservation but is also practical for large scale institutions. It now underpins advice and guidance given to the UK archive sector on digital preservation. (11 pages).



Community Owned digital Preservation Tool Registry COPTR

http://coptr.digipres.org/Main_Page

COPTR describes tools useful for long term digital preservation and acts primarily as a finding and evaluation tool to help practitioners find the tools they need to preserve digital data. COPTR aims to collate the knowledge of the digital preservation community on preservation tools in one place. It was initially populated with data from registries run by the COPTR partner organisations, including those maintained by the Digital Curation Centre, the Digital Curation Exchange, National Digital Stewardship Alliance, the Open Preservation Foundation, and Preserving digital Objects With Restricted Resources project (POWRR). COPTR captures basic, factual details about a tool, what it does, how to find more information (relevant URLs) and references to user experiences with the tool. The scope is a broad interpretation of the term "digital preservation". In other words, if a tool is useful in performing a digital preservation function such as those described in the OAIS model or the DCC lifecycle model, then it's within scope of this registry



DPC Getting Started in Digital Preservation Workshops

<http://www.dpconline.org/events>

The DPC Getting Started in Digital Preservation workshops are events designed to raise awareness of digital preservation issues, increase involvement with digital preservation activities and sign-post the support and resources available to help you on your way. They provide an introduction to digital preservation, build an understanding of the risks to digital materials, include practical sessions to help you apply digital preservation planning and tools, and feature speakers sharing their own experience of putting digital preservation into practice. You can find details of forthcoming workshops and the programmes and speaker presentations at previous workshops on the DPC events page.

Digital Preservation Management: Implementing Short-Term Strategies for Long-Term Solutions

<http://www.dpworkshop.org/>

An excellent free online tutorial that introduces you to the basic tenets of digital preservation. It is particularly geared toward librarians, archivists, curators, managers, and technical specialists. It includes definitions, key concepts, practical advice, exercises, and up-to-date references. The tutorial is available in English, French, and Italian.

Canadian Heritage Information Network (CHIN) Digital Preservation Toolkit

http://www.rcip-chin.gc.ca/carrefour-du-savoir-knowledge-exchange/outils_preservation_numerique-digital_preservation_toolkit-eng.jsp

CHIN has released a suite of documents to identify digital material found in museums, the potential risk and impact of lost material, and how to get started in the development of Preservation Policies, Plans and Procedures. The toolkit includes a Digital Preservation Inventory Template, Digital Preservation Policy Framework Development Guideline, Decision Trees, and a Digital Preservation Plan Framework.

Digital Preservation 101, or, How to Keep Bits for Centuries

http://scholar.harvard.edu/jcs/presentations/dhttp://handbook.dpconline.org/administrator/index.php?option=com_content&view=article&layout=edit&id=86igital-preservation-101-or-how-keep-bits-centuries

This 2015 presentation by Julie Swierczek Digital Asset Manager and Digital Archivist at Harvard University Art Museums is a good advocacy for and explanation of, digital preservation to other non-specialist institutional colleagues including "why archivists cry themselves to sleep at night when the general public conflates archives with backup copies of data" (142 slides but many are images with good slide notes that make this easily understandable).

The National Archives Digital Continuity Guidance

<http://www.nationalarchives.gov.uk/information-management/manage-information/policy-process/digital-continuity/step-by-step-guidance/step-2/>

This guidance takes you through the process of creating an information asset register, and includes a template in Excel spreadsheet format. The register can be useful for Records Managers/Information Managers as a model which they can demonstrate aligns with business risk management.



Risk Management for Digital Preservation

<https://vimeo.com/171082277>

From a series of video covering topics from the 'Getting Started in Digital Preservation' roadshows, this video provides a brief introduction to the use of risk management for Digital Preservation.

Case studies



Bishopsgate library case study

http://wiki.dpconline.org/index.php?title=Bishopsgate_library_case_study

A collections audit and business case focused on taking the first steps of digital preservation at the Bishopsgate Institute Library. (28 pages).

Starting Small: Practical First Steps in Digital Preservation

<http://www.slideshare.net/hakbailey/starting-small-practical-first-steps-in-digital-preservation-13385434>

One example of how digital preservation principles can be added to the collections management activities of a small institution (Dartmouth College USA from 2010–2012), without needing a lot of additional resources. (26 slides).

DPC case note: West Yorkshire Archive Service accepts a digital collection

http://www.dpconline.org/component/docman/doc_download/511-casenotemlawyas.pdf

In this Jisc-funded case study staff from West Yorkshire Archives Service report on their experience in taking their first large digital archive. This made them confront new problems and new ways of working, they conclude that "If we try we may fail; if we don't try we will certainly fail". October 2010 (4 pages).

DPC case note: Glasgow Museum takes first steps in turning an oral history headache into an opportunity

http://www.dpconline.org/component/docman/doc_download/502-casenoteglasgowmuseums.pdf

This Jisc-funded case study examines how Glasgow Museums' took some simple steps in addressing digital preservation and created short and long term opportunities. Activities such as creating an inventory, assessing significance and promoting access provide the basis for building confidence to manage the wider challenges, and can bring early rewards if properly embedded within the mission of an organization. September 2010 (4 pages).

Digital Preservation Planning Case Study

http://www.dpconline.org/component/docman/doc_download/863-2013-may-getting-started-london-planning-case-study-ed-fay

A set of DPC Getting Started in Digital Preservation workshop presentation slides by Ed Fay from May 2013. An excellent concise overview of planning for digital preservation and how to approach it . (20 slides).

References

Dollar, C.M. and Ashley, L.J., 2014. *Assessing Digital Preservation Capability Using a Maturity Model Process Improvement Approach*. Available:

<http://static1.squarespace.com/static/52ebbb45e4b06f07f8bb62bd/t/53559340e4b058b6b2212d98/1398117184845/DPCMM+White+Paper+Revised+April+2014.pdf>

NDSA , 2013. *The NDSA Levels of Digital Preservation: An Explanation and Uses, version 1 2013*.

National Digital Stewardship Alliance. Available:

http://www.digitalpreservation.gov/ndsa/working_groups/documents/NDSA_Levels_Archiving_2013.pdf

Digital Preservation Handbook

Content Specific Preservation



Illustrations by Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Who is it for?

Operational managers (DigCurV Manager Lens) and staff (DigCurV Practitioner Lens) in repositories, publishers and other data creators, third party service providers.

Assumed level of knowledge

Novice to Intermediate

Purpose

- To provide a bridge to and achieve synergies with, reports in the DPC Tech Watch Series. The reports provide advanced level "deep dives" in specific areas of content preservation (e.g. email) that can be cited or to source case studies in the Handbook.
- To be developed for ease of maintenance, cost-efficiency, and sustainability in the long-term by the DPC via updates and additions to the Tech Watch series.
- To provide a brief overview and case studies, suitable for novice or intermediate level users, of digital preservation issues for specific content types covered by DPC Technology Watch Reports. Currently three content types are available: e-journals, moving picture and sound, and web-archiving. We hope to add more at a later date.

Gold sponsor



Silver sponsors



Bronze sponsors



Reusing this information

You may re-use this material in English (not including logos) with required acknowledgements free of charge in any format or medium. See [How to use the Handbook](#) for full details of licences and acknowledgements for re-use.

For permission for translation into other languages email: handbook@dpconline.org

Please use this form of citation for the Handbook: Digital Preservation Handbook, 2nd Edition, <http://handbook.dpconline.org/>, Digital Preservation Coalition © 2015.

Contents

e-Journals	4
Case study 1: the e-journal or its past issues are no longer available from the publisher	6
Case study 2: library e-Journals, perpetual access, and de-accessioning print.....	7
Resources	8
References.....	9
Moving pictures and sound.....	10
Case study 1: The Open University (OU) Access to video assets project	13
Case study 2: British Library Archival sound recordings project	13
Case Study 3: Imperial War Museum PSRE project.....	13
Case Study 4: British University Film and Video Council Newsfilm Online project	13
Case Study 5: BFI and Regional Film Archives Screen Heritage UK (SHUK) project	14
Resources	15
Further case studies	17
References.....	18
Web-archiving	18
Case study 1: The UK Web Archive	21
Case study 2: The Internet Memory Foundation	22
Case study 3: The Coca-Cola web archive	23
Resources	24
Further case studies	27
References.....	28

e-Journals



Illustration by Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Overview

This case study provides a brief novice to intermediate level overview for e-journal preservation summarised from the DPC Technology Watch Report Preservation, Trust and Continuing Access for e-Journals with updates and additions by the author. Two "mini case studies" are included together with short summaries of major services and solutions. The report itself is recommended to readers who need a more advanced level briefing on the topic and practice. It covers a wider range of issues and practice in greater depth with extensive further reading and advice ([Beagrie, 2013](#)).

Introduction

Digital Preservation and trust in having continuing future access to digital content have become increasingly important for research libraries as published journals and articles have shifted from print to electronic formats. Traditional publishing business models and relationships have also undergone major transformations as a result of that shift.

Among many significant changes there has been a move from libraries purchasing and physically holding (and preserving) a paper journal locally (with multiple redundancy of copies between libraries), to renting (licensing) remote access to an electronic journal held on publishers' platforms that are often based internationally in other jurisdictions.

In parallel, there has been a growing open-access movement for e-journal articles that seeks to remove the subscription charges for access. Subscription journals, open-access journals and hybrids of the two (either a mixture of open-access and subscription articles in a journal or a 'moving wall' to open access after a fixed period of time) provide a complex landscape for the preservation of, and long-term access to, e-journals.

This e-journal landscape continues to evolve as e-publishing itself begins to shift from static to dynamic content, and the importance of data and supplementary material linked to articles increases in major disciplines.

All these changes in turn have made preservation of e-journals more demanding, more international and dependent on others, and brought issues of trust to the fore. Trust in this context is not solely of technology for preservation, but negotiating rights (and retaining a record of them for future use), and having transparent information on what is being archived, how it is preserved, and how and when it can be accessed.

This makes e-journals one of the most dynamic and challenging areas of digital preservation, particularly in terms of business models and trust mechanisms for shared or out-sourced preservation services.

Services and solutions

It is important to understand the significant implications for preservation and access of the different requirements (and terminology) that apply for e-journals: in particular the distinction between continuing access and long term preservation, as these differences lead to different types of service for e-journal archiving.

- **Continuing access** (sometimes also called post-cancellation or perpetual access) applies only to subscription journals and securing long-term access for their subscribers;
- **Long-term preservation** applies to both open and subscribed content.

The main preservation and continuing access services and solutions available for e-journals are as follows:

Keepers Registry
The Keepers Registry is a Jisc service to provide easily accessible information about inclusion of e-journals in preservation services and to highlight those e-journals for which no archiving arrangements exist. EDINA, a national data centre based at the University of Edinburgh, has developed the service along with its partner in the project, the ISSN International Centre in Paris
Legal and voluntary deposit in copyright libraries
The role of a national library is to ensure that the published heritage of its country is preserved and made accessible. In many countries legal deposit is an important vehicle for achieving this is. There is a global trend towards extending legal deposit from the print environment to cover e-journals and other electronic publications. Legal deposit legislation (or similar voluntary deposit arrangements) normally involves those subscription e-journals considered part of the national published heritage of that country. To protect the commercial interests of the publisher it also restricts off-site access to preserved electronic material for a substantial period of time. Typically this means a national legal deposit collection does not cover the international range of subscription e-journals licensed by other libraries and their users, and does not meet their requirements for ‘perpetual access’ rights.
CLOCKSS
CLOCKSS (Controlled LOCKSS) is a not-for-profit collaboration between libraries and publishers. It is a dark archive based on the LOCKSS software (see section below on LOCKSS) in which a limited number of libraries take on an archival role on behalf of a broader community. It provides insurance to libraries that the e-journal and other content they have subscribed to will be preserved for the long term. It is described as a ‘private LOCKSS network’.

KB e-Depot

The Koninklijke Bibliotheek (KB) is the national library of the Netherlands and operates the e-Depot. It has taken the policy decision to archive journals that are within its national mandate and additionally a range of e-journals (including open-access titles in the Directory of Open-Access Journals) published beyond its borders. The e-Depot does not currently provide for post-cancellation continuing access by licensees of the content. Generally, end-user access is restricted to on-site perusal at the KB for reasons of private research only and online access is denied. However, full online access is granted to publications by open-access publishers.

LOCKSS

LOCKSS (Lots of Copies Keep Stuff Safe) provides libraries with open-source tools and support so they can take local custody of a wide variety of materials, including subscription and open-access scholarly assets (books, journals, etc.). Readers access LOCKSS preserved content whenever (and for whatever reason) the material cannot be viewed on the publisher's (or intermediary's) servers. The highly distributed nature of this approach aims to ensure that there is sufficient replication to safeguard content despite any potential disasters which might befall individual LOCKSS institutions.

Portico

Portico is designed specifically as a third-party service for scholarly literature published in electronic form and provides three specific preservation services for e-journals, e-books and digitized historical collections respectively. It provides insurance to libraries that the e-journal and other content they have subscribed to will be preserved for the long term. Portico only provides access to the e-journals they have preserved after specified 'trigger events'. In addition, if a publisher has designated Portico as such, it can also serve as a potential mechanism for post-cancellation access.

Consortial hosting

A small number of regional consortia also organize and provide their own hosting services for access and preservation of e-journals. Notable examples are OhioLink, operated by the Ohio Library and Information Network, and the Scholars Portal, operated by the Ontario Council of University Libraries.

Case study 1: the e-journal or its past issues are no longer available from the publisher

This is a highly likely scenario as publishers merge or change their business models, as larger publishers review and adjust their portfolio of titles, or as learned societies move publication contracts for their journals from one publisher to another. Journal titles are also sometimes traded between publishers, which may mean that access to past issues is no longer supported by the previous owner.

The UKSG Transfer Code of Practice initiative has produced a Code of Practice aimed at easing the problems created when journal titles move between publishers. Of relevance are the following paragraphs contained in version 3 of the code ([UKSG, 2014](#)):

The transferring publisher will alert the receiving publisher to all existing preservation arrangements for the journal.

The transferring publisher must ensure continued access to its subscribers where it has granted perpetual access rights, even if the transferring publisher will cease to host the online version of the journal after the effective transfer date. Either the transferring or the receiving publisher, or both, could fulfill perpetual access obligations. The Code intentionally does not specify the means for achieving such access, but places on the transferring publisher the responsibility for ensuring that subscribers to whom it has granted perpetual access rights will continue to have access post-transfer.

The transferring publisher will use reasonable efforts to communicate journal transfer information where perpetual access rights were granted as part of a licensing agreement/Big Deal, unless archival rights will remain with the transferring publisher.

Subscribers that have been granted perpetual access rights to previously published content with the authority of the journal owner must have those rights honoured. Either the transferring or the receiving publisher, or both, could fulfil perpetual access obligations.

The receiving publisher will continue the existing, or equivalent, preservation arrangements for the journal after the effective transfer date. The receiving publisher will not remove content that was previously deposited in preserving archive(s), even if the receiving publisher will not be continuing to deposit content in the archive(s).

The decision of the publisher Sage to no longer offer its publication *Graft* provided a real-life example of triggered access from three archiving solutions – Portico, KB e-Depot, and CLOCKSS. In this case all were able to continue to offer access to the issues they held, either as open access (CLOCKSS and KB e-Depot) or else as a service to members (Portico). While it cannot be guaranteed that the archive will include all back issues of the title (as with *Graft*), participation in an archiving solution which covers at least some issues will significantly reduce the risk of disruption to continuity of service.

Case study 2: library e-Journals, perpetual access, and de-accessioning print

This case-study was first published by Jisc as part of work funded in its digital preservation programme and was incorporated into the Tech Watch Report. It has been adapted for use in the Handbook.

The case study differs from others in illustrating a few of the issues in realizing some of the potential cost savings from e-journals, particularly space savings. Increasingly, academic libraries are investing heavily in e-journals which duplicate their print back-runs. For libraries facing acute pressures on space, one solution to their problem is to dispose of or relegate print back-runs which overlap with their electronic holdings.

The case study focuses on work at Imperial College London Library in providing a database and toolkit for staff making such de-selection decisions ([Cooper and Norris, 2007](#)). Imperial established three criteria to determine the sustainability of their e-journals for de-accessioning of print. Their electronic access was classified as sustainable when at least one of the following applied:

- Imperial had perpetual access rights to the content, via the web. Imperial's perpetual access rights were nowhere near as comprehensive as they would have wished; they estimated that less than 50% of their content was covered. In addition, some of their licences specified an unsuitable delivery method for post-termination access. As they were no longer supporting networked CD-ROMs and did not have the resources to mount journal content locally, they considered a journal sustainable only if perpetual access is provided via the web.
- The journal was permanently open access for all years or certain years. Hybrid open-access journals were not included in this category, as the project was not interested in sustainability

at the article level. Finding open-access journals which fulfilled their criteria proved harder than anticipated. The main stumbling block was their need for assurance on the permanency of open access. Although the Bethesda and Berlin Declarations on Open Access include perpetual access in their definitions, Imperial discovered that not all 'open-access journals' met this criterion of permanency.

- The content was in one of Imperial's trusted services such as JSTOR, the ACM digital archive or a Jisc-funded archive. Imperial noted that of their three sustainability criteria, this one, covering services that did not offer perpetual access rights, was the hardest to pin down. The services falling into this category all shared two characteristics: the first was a good track record of stability, i.e., they had demonstrated continuity of titles from one year to another for as long as they had subscribed; the second was a history of and reputation for, affordability and value for money.

Twenty-one months into the project Imperial had identified 700 shelf-metres of sustainable stock for disposal from one site, and planned to rollout the de-selection exercise to other sites. Although it was still early days, they felt their sustainability criteria seemed to be working. The only sustainable content that they had lost was four journals from the same publisher, and they were in the process of challenging that loss. This proved to be an added benefit of the entitlements database they had created for the project; without it they would not have been aware that content over which they had perpetual access rights had been lost.

Conclusions

Continuing access and preservation of e-journals has involved initiatives in organizing multi-institutional collaboration, developing third-party services, and establishing trust in long-term access and preservation between different stakeholders. The issues it has had to address go well beyond technology, and legal, economic and service developments are equally critical to its success. Many challenges remain in e-journal archiving, but there have been significant successes and lessons learnt of interest to the wider digital preservation community as well as to libraries and publishers.

Resources



Preservation, Trust and Continuing Access for e-Journals, DPC Technology Watch Report 13-04 September 2013

<http://dx.doi.org/10.7207/twr13-04>

This report discusses current developments and issues which libraries, publishers, intermediaries and service providers are facing in the area of digital preservation, trust and continuing access for e-journals. It also includes generic lessons and recommendations on outsourcing and trust learnt in this field of interest to the wider digital preservation community. It is not solely focused on technology, and covers relevant legal, economic and service issues (43 pages).

To bin or not to bin? Deselecting print back-runs available electronically at Imperial College London Library

<https://spiral.imperial.ac.uk/handle/10044/1/503>

Increasingly, academic libraries are investing heavily in e-journals which duplicate their print back-runs. For libraries facing acute pressures on space, one solution to their problem is to dispose of or relegate print back-runs which overlap with their electronic holdings. This 2007 article by R Cooper and D Norris describes work at Imperial College London Library to provide a tool-kit for staff making such de-selection decisions.

UKSG, 2014 Transfer Code of Practice: Version 3.0 March 2014

<http://www.uksg.org/Transfer/Code>

The Transfer Code of Practice promotes a set of standards that apply whenever a journal is transferred from one publisher or publishing platform to another. Publishers who publicly sign up to the Code and apply it in practice are considered 'Transfer compliant'. As a voluntary best practices code for industry participants, the Transfer Code of Practice does not supplant contractual terms, intellectual property rights or the competitive marketplace between publishers.



CLOCKSS

<http://www.clockss.org>

KB e-Depot

<http://www.kb.nl/en/organisation/research-expertise/long-term-usability-of-digital-resources/information-for-international-publishers>

LOCKSS

<http://www.lockss.org>

Portico

<http://www.portico.org>

Ohio Link

<http://www.ohiolink.edu>

Scholars Portal

<http://www.ocul.on.ca/node/135>

Keepers Registry

<http://thekeepers.org>

References



Beagrie, N., 2013. Preservation, Trust and Continuing Access for e-Journals *DPC Technology Watch Report* 13-04 September 2013. Available: <http://dx.doi.org/10.7207/twr13-04>

Cooper, R. and Norris, D., 2007. *To bin or not to bin? Deselecting print back-runs available electronically at Imperial College London Library*, *Serials* 20 (3), 208–214. Available: <https://spiral.imperial.ac.uk/handle/10044/1/503>

UKSG, 2014. *Transfer Code of Practice: Version 3.0* March 2014. Available: <http://www.uksg.org/Transfer/Code>

Moving pictures and sound



Illustration by Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Overview

This case study provides a brief novice to intermediate level overview summarised from the DPC Technology Watch Report on Preserving Moving Picture and Sound. Five "mini case studies" of UK collections that have run preservation and access projects for sound and moving image content are included. The report itself provides a "deep dive" discussing a wider range of issues and practice in greater depth with extensive further reading and advice ([Wright, 2012](#)). It is recommended to readers who need a more advanced level briefing on the topic and practice.

Introduction

The audiovisual domain is unique in that digitization is routinely critical to preservation. Audiovisual digitization for preservation is so pervasive that the two words have come to be used interchangeably. Audio and video need digitization for the very survival of their content, owing to the obsolescence of playback equipment and decay and damage of physical items, whether analogue or digital. The basic technology issue for collections of moving images and sound is the necessity to digitize all content currently sitting on shelves. Film on shelves can be conserved (unless it is already deteriorating), but still needs digitization to provide access.

A vital issue in preservation is access: motivation and funding for digitization purely for preservation purposes is difficult, if not impossible. There is great public, institutional and educational interest in the audiovisual record of the twentieth century. Creating access to that record is the key to obtaining the support needed for the digitization and preservation of the content.

The landscape for 'moving pictures and sound' is complicated: physically, there are large differences between audio, video and film recordings. The formats and record/playback equipment are completely separate; the digitization procedures are different; the digital files have different wrapper formats and metadata (with some overlaps); and the storage requirements differ, with video taking roughly 100 times as much storage per second of material as does audio, and high resolution digital film taking roughly 10 times more storage than video.

In addition culturally and economically, there are significant preservation and curation differences between collections from:

- **commercial media industries** – music, cinema and commercial broadcasting where preservation needs a commercial justification, a business case;
- **public bodies** – public service broadcasting, academic collections and heritage institutions such as national museums, libraries and film institutes where preservation needs a cultural heritage justification, though increasingly this sector also needs a business case;
- **technical areas** such as medicine, geology and surveillance, where recordings of images or of seismic events are raw data, kept as medical records or for reprocessing; and
- **other** – a wide range of independent collections, ranging from individual efforts to material gathered by non-profit specialist institutions (for example, steam engine clubs or ethnological research) that do not fall into any of the above categories, though their material may eventually end up being donated to a public collection.

Within the landscape is a range of technologies including engineering, computing, Internet technology, archiving, media management, museum collections management, curation, preservation, access, knowledge management and resource discovery.

Technical challenges

Audiovisual recordings are surrogate reality. The technology allows the listener and viewer to get a sensation of what a situation sounded and looked like, but the technology actually only captures the sequence of light patterns or sound pressures acting on the recording instrument (camera, microphone). These patterns (for film) and signals (for video and audio) are more like data than like artefacts. The preservation requirement is not to keep the original recording media, but to keep the data, the information, recovered from that media.

A key technology issue is moving digital content from carriers (such as CD and DVD, digital videotape, DAT and minidisc) into files. This digital to digital 'ripping' of content is an area of digital preservation unique to the audiovisual world, and has unsolved problems of control of errors in the ripping and transfer process.

The final technology area is digital preservation of the content within the files that result from digitization or ripping, and the files that are born digital. While much of this preservation has problems and solutions in common with other content, there is a specific problem of preserving the quality of the digitized signal that is again unique to audiovisual content. Managing quality through cycles of lossy encoding, decoding and reformatting is one major digital preservation challenge for audiovisual files. The other issue is managing embedded metadata.

For three decades for audio, and for at least two decades for video, archives have been digitizing their analogue content for preservation and access. The problem areas are:

- successful playback of the originals, in order to get an optimal signal to digitize;

- standards: what compression level, encoding method and file format to use; and
- efficiency: digitizing the existing analogue materials fast enough and economically enough to cope with the size and urgency of the problem.

Stages in sound and moving image digital preservation

For sound and moving image preservation, the following stages in the overall process need to be kept clear:

- **signal:** the audio from a microphone, the video signal coming out of a video camera. These signals have physical properties (bandwidth; dynamic range) that can be defined and measured. The quality of a recording and the success or failure of any process of copying, digitization or preservation can be reduced (in large part) to how well that process maintains these two physical properties of the original signal;
- **recording of a signal onto a carrier** (also called support, physical medium or recording format). For a century, the methods of capturing a signal were tied to the carrier of the signal: a wax cylinder, film reel or videotape. Digital technology produces recordings that are independent of carriers. Carrier independence is liberation: discs, tapes and films deteriorate or get damaged. Born digital recordings are liberated from these carrier-based problems, leading to a desire to liberate analogue recordings by digitization;
- **digitization:** analogue recordings can be played back and recorded onto a new carrier, or digitized and so released from carrier dependence. Digitization has to ensure that the digital version has the same bandwidth and dynamic range as the original, to capture the original quality; and
- **digital preservation of the digital representation of a signal**, meaning preserving the numbers, but also preserving the technology needed to decode (render) the numbers. Audiovisual content has a particular problem. The coding of the signal can be a compromise, not actually capturing the full signal, but instead losing some of it (lossy encoding) to get a more compact representation, thus reducing storage and transmission costs. Unfortunately coders/decoders (codecs) go out of use, and are replaced by newer technology. The file format holding the coded signal, the wrapper, is also subject to obsolescence. The failure and obsolescence of storage technology and the obsolescence of encode/decode methods and wrapper formats are major digital preservation problems for audiovisual content.

Access and rights

Sound and moving picture content arising from cinema, broadcasting and the commercial music industry is constrained by rights issues. Music has copyright protection for the composer and for the physical object containing a performance (so-called magnetic copyright). Cinema productions are protected, and music used in a film retains its separate protections. Broadcasting is even more complicated, as all the parties involved in a production may have rights in future exploitation subsequent to the one or two transmissions that were specified in typical contracts. These rights are seen as protection by rights holders, but are also seen as restrictions on access. The situation for a public broadcaster is particularly difficult. The public invariably feel that any production by a public broadcaster has already been paid for by them, is already publicly owned and should be available for public access. Unfortunately that understandable feeling is not the same as the legal definition governing when a work enters the public domain (usually determined by expiry dates on copyright and other rights).

Case study 1: The Open University (OU) Access to video assets project

This is an access and re-use project. The focus is to digitize (where necessary) audiovisual assets previously created by the OU, and place them in an asset management system so that current OU teaching and other activity can find and use these assets. Preservation is a by-product of the project rather than an end in itself. This project provides an important example of combining preservation of content with use of content, something of value to the institution in order to obtain a budget and deliver a benefit. The project was presented at the DPC Briefing Day 'Preserving Digital Sound and Vision'. The project digitized 1,200 videotapes and films, and placed the results in a Fedora digital repository. Also, 145,000 pages of documentation were digitized, providing the overall educational framework around the 1,200 items, giving them context and enhancing their ability to be re-used. The user interface provides granularity and time-based navigation. Overall this project is an outstanding example of best practice.

Case study 2: British Library Archival sound recordings project

This is a JISC-supported preservation and educational access project that ran (in its initial phase) from 2004 to 2006. A second phase added further material. Nearly 50,000 recordings of speech, music and sounds of 'human and natural environments' were digitized and placed online. The online catalogue is open to all and licensed UK further or higher education institutions can also listen to the audio. Anyone can listen to 2,000 of the items (or any of them by attending the British Library reading room in London). The differences in access between educational institutions and the general public reflects the overall issue of rights as the one remaining constraint on open access to audiovisual materials in public institutions.

Case Study 3: Imperial War Museum PSRE project

The Imperial War Museum has one of the UK's major film collections. It has been collecting film since its founding in 1919, beginning with footage from the Great War that led to the institution's founding. The Public Sector Research Exploitation (PSRE) fund made an award of nearly £1 million for cataloguing, digitization and online access (to the catalogue and the footage). The project ran from 2006 to 2009 and is of particular interest in that it is specifically aimed at commercial exploitation of a collection, and at sustainable business models around digitization and web access. The result is a website (<http://film.iwmcollections.org.uk/>) where anyone can view content in low quality; pull documents, stills and key frames into a lightbox; and fill a shopping basket to then purchase content.

Case Study 4: British University Film and Video Council Newsfilm Online project

This is another project with JISC sponsorship. For four decades to 1960 newsreels shown in cinemas were the main way for the general public to see moving images of current events. The initial project ran from 2004 to 2008. The results are available through a website which, as for the BL Archival Sound Recordings project, has full functionality for registered universities and colleges. The general public can see the full catalogue and can see a single key frame for each item. Since the original phase of the project, the content has been augmented by ITN/Reuters news covering the events from decades after the decline of newsreels. Newsreel items are short: the initial project provided 3,000 hours of content, but that represented 60,000 items. In addition, as with the Open University project, documentation was also placed online for context and to support search and retrieval: 450,000 pages of bulletin scripts.

Case Study 5: BFI and Regional Film Archives Screen Heritage UK (SHUK) project

SHUK is a large (£22.8 million) and complex project (involving 12 regional film archives in addition to the BFI). The project was complicated by changes in the structure and funding of the BFI, as well as a change of government and a raft of other issues. Nevertheless the project has produced major achievements:

- conservation, not digitization: construction of a £6-million vault for film conservation;
- digitization: film scanning and digital storage equipment for the regional film archives;
- access: online catalogues of regional film archive content, available to the general public.

SHUK launched on 5 September 2011 with a BBC BFI joint production, *The Reel History of Britain* ([SHUK, 2011](#)).

Conclusions

The basic technology issue for collections of moving images and sound is the necessity for digitization of all content that is currently sitting on shelves. Audio and video need digitization for their very survival, owing to obsolescence and decay of physical items, whether analogue or digital. Film on shelves can be conserved (unless it is already deteriorating) but needs digitization for access.

Playback for preservation-quality digitization implies the need for optimal recovery of the original quality, which requires professional equipment and experience. The major technical obstacle is that, for many physical formats, the needed equipment is largely obsolete, meaning that parts and repairs and skilled operators are in increasingly short supply. The urgent recommendation is, do not wait! Audiovisual holdings need to be documented and made part of a preservation plan.

The situation for sound heritage is clear. The digitization standards, encoding, wrapper and metadata are all agreed and well documented in IASA TC-04 ([IASA, 2009](#)). Uncompressed audio in the Broadcast Wave Format (BWF) wrapper is widely used and well supported. There is no reason for the basic encoding to ever be changed, though the BWF wrapper may eventually become obsolete. The only significant problem is the failure of some standard audio applications to handle embedded BWF metadata correctly ([ARSC, 2011](#)). All archives need to be aware of the risk of loss of embedded metadata. The situation for video is complex, but there is a PrestoSpace roadmap for guiding choices on the digitization of various legacy formats. There is advice from the PrestoCentre and from JISC Digital Media on the digital preservation of the resultant files. A big challenge is a registry of applications that work properly on embedded video metadata, where the diversity is huge. There is no single agreed wrapper, metadata standard or even encoding standard, and the change from standard definition to high definition brings a new set of applications, wrappers and encodings.

There is emerging technology that can improve audio (capture of the bias tone and consequent removal of temporal variation) and video transfers (direct digitization of the RF signal from the read head), which could be useful in those cases where current technology fails. So the recommendation is not to wait until such technology is further advanced and more widely available. If there are playback problems that cannot be resolved, the original audio or video format should be kept so that such advanced technology can be applied in the future.

Quality checking of the results of digitization remains an issue for video. There is a need for effective integration of signal processing technology with human checking in order to produce a really efficient method of quality control within a preservation factory approach. Quality checking is equally relevant to digital preservation – any changes or migrations due to digital obsolescence need to be checked for preservation of signal quality. Again, a purely manual approach does not scale (to the tens of millions of hours of audiovisual content in European collections), while purely algorithmic substitutes for 'looking and listening' have never been completely successful and remain an area where further research is needed.

Resources



Wright, R., 2012. Preserving Moving Pictures and Sound DPC Technology Watch Report 12-01 March 2012

<http://dx.doi.org/10.7207/twr12-02>

This report is for anyone with responsibility for collections of sound or moving image content and an interest in preservation of that content. New content is born digital, analogue audio and video need digitization to survive and film requires digitization for access. Consequently, digital preservation will be relevant over time to all these areas. The report concentrates on digitization, encoding, file formats and wrappers, use of compression, obsolescence and what to do about the particular digital preservation problems of sound and moving images (*33 pages*).

SHUK, 2011. Screen Heritage UK Marks new Era for Britain's Film Archives

<http://www.bfi.org.uk/sites/bfi.org.uk/files/downloads/bfi-press-release-screen-heritage-uk-marks-a-new-era-for-britains-film-archives-2011-09-01.pdf>

BFI Press release. 8 pages

IASA 2009 IASA TC-04, Guidelines on the Production and Preservation of Digital Audio Objects (IASA-TC 04 Second edition 2009) Canberra, IASA.

<http://www.iasa-web.org/audio-preservation-tc04>

This is the standard guide to digitization of audio, and the sections on metadata and digital storage are of value to all forms of digital media.

Casey, M. and Gordon, B., 2007. Best Practices for Audio Preservation. Bloomington, Indiana University Bloomington.

<http://www.dlib.indiana.edu/projects/sounddirections/papersPresent/>

Another audio resource (that also includes a range of digitization software tools) comes from the Sound Directions project of Harvard and Indiana Universities: much is also relevant to video digitization. (*160 pages*)

Digital Preservation Coalition Briefing day on Preserving Digital Sound and Vision, April 2011

<http://www.dpconline.org/events/details/27-SoundAndVision?xref=26>

This DPC briefing day in April 2011 provided a forum to review and debate the latest development in the preservation of digital sound and vision. Seven presentations (including the Open University) are linked from the programme and available to download.

ARSC Technical Committee, 2011. Study of Embedded Metadata Support in Audio Recording Software. Association of Recorded Sound Collections.

http://www.arsc-audio.org/pdf/ARSC_TC_MD_Study.pdf

A study of support for embedded metadata within and across a variety of audio recording software applications. The findings raise serious concerns, particularly for the archiving and preservation communities who rely on embedded metadata for interpretation and management of digital files representing preserved content into the future. (21 pages)



AVPreserve

<http://www.avpreserve.com/>

US based media and information management consulting firm. Its website provides a range of resources for AV preservation.

BUFVC NewsFilm online Project

<http://www.webarchive.org.uk/wayback/archive/20140614061518/http://www.jisc.ac.uk/whatwedo/programmes/digitisation/bufvc.aspx>

British Film Institute

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

the British Film Institute can advise on film and also on video – they hold a lot of video, and have a Curator for Television. Its remit is collection and preservation of film and television, and technical advice.

British Library Sound Archive

<http://www.bl.uk/nsa>

General technical advice on audio preservation is available from the British Library Sound Archive. Its remit is collection and preservation of all forms of audio, and technical advice.

Film Archives UK

<http://filmarchives.org.uk>

Collection and preservation of general audiovisual content of regional significance in the UK

JISC Digital Media

<http://www.jiscdigitalmedia.ac.uk>

Advice and training on still images, moving images and sound. This includes their InfoKits for Digital File Formats, Digitisation funding and sustainability, and High Level Digitisation Guide for Audiovisual Resources.

PrestoCentre

<http://www.prestocentre.eu>

Website provides audiovisual information, resources and advice. Access to most resources on the website requires a member subscription but a number are available to non-numbers.

The Preservation Guide Wiki

<http://preservationguide.co.uk/RDWiki/>

This audiovisual preservation guide was created for PrestoSpace and the BBC in May 2006. The site is now maintained as part of The Preservation Guide Consultancy. The wiki is in the public domain under a creative commons licence.

Sustaining Consistent Video Presentation

<http://www.tate.org.uk/research/publications/sustaining-consistent-video-presentation>

This technical paper addresses approaches to identifying and mitigating risks associated with sustaining the consistent presentation of digital video files. Originating from two multi-partnered research projects – Pericles and Presto4U – the paper was commissioned by Tate Research and is intended for those who are actively engaged with the preservation of digital video.



JISC 2009 - Archival Sound Recordings Showreel

<https://www.youtube.com/watch?v=KPy9ZqWEHog>

Engaging short video on British Library archival sound recordings project published on 22 Jun 2009. (6 mins 11 secs).

Further case studies



Podcasts in the Archives: Archiving Podcasting Content at the University of Michigan

<http://files.archivists.org/pubs/CampusCaseStudies/CASE12.pdf>

In this Society of American Archivists campus case study Alexis. A. Antracoli, University of Michigan, examines the challenges involved in developing best practices and workflows for archiving and preserving podcasting content. One major issue involved establishing standards of practice for ingest, storage, and access, especially the generation and storage of appropriate descriptive, technical, and

preservation metadata. Another challenge centered around developing the necessary technological infrastructure to support an Open Archives Information System (OAIS)-compliant system. 2010. (14 pages).

References

ARSC Technical Committee, 2011. *Study of Embedded Metadata Support in Audio Recording Software*. Association of Recorded Sound Collections. Available: http://www.arsc-audio.org/pdf/ARSC_TC_MD_Study.pdf

IASA, 2009. *IASA TC-04, Guidelines on the Production and Preservation of Digital Audio Objects*, IASA-TC 04 Second edition 2009, Canberra, IASA. Available: <http://www.iasa-web.org/audio-preservation-tc04>

SHUK, 2011. *Screen Heritage UK Marks new Era for Britain's Film Archives*. Available: <http://www.bfi.org.uk/sites/bfi.org.uk/files/downloads/bfi-press-release-screen-heritage-uk-marks-a-new-era-for-britains-film-archives-2011-09-01.pdf>

Wright, R., 2012. *Preserving Moving Pictures and Sound DPC Technology Watch Report 12-01 March 2012*. Available: <http://dx.doi.org/10.7207/twr12-02>

Web-archiving



Illustration by Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Overview

This case study provides a brief novice to intermediate level overview summarised from the DPC Technology Watch Report on Web-Archiving. Three "mini case studies" are included illustrate the different operational contexts, drivers, and solutions that can be implemented. The report itself provides a "deep dive" discussing a wider range of issues and practice in greater depth with extensive further reading and advice ([Pennock, 2013](#)). It is recommended to readers who need a more advanced level briefing on the topic and practice.

Introduction

The World Wide Web is a unique information resource of massive scale, used globally. Much of its content will likely have value not just to the current generation but also to future generations. Yet the lasting legacy of the web is at risk, threatened in part by the very speed at which it has become a success. Content is lost at an alarming rate, risking not just our digital cultural memory but also organizational accountability. In recognition of this, a number of cultural heritage and academic institutions, non-profit organizations and private businesses have explored the issues involved and lead or contribute to development of technical solutions for web archiving.

Services and Solutions

Business needs and available resources are fundamental considerations when selecting appropriate web archiving tools and/or services. Other related issues must also be considered: organizations considering web archiving to meet regulatory requirements must, for example, consider associated issues such as authenticity and integrity, recordkeeping and quality assurance. All organizations will need to consider the issue of selection (i.e. which websites to archive), a seemingly straightforward task which is complicated by the complex inter-relationships shared by most websites that make it difficult to set boundaries. Other issues include managing malware, minimizing duplication of resources, temporal coherence of sites and long-term preservation or sustainability of resources. International collaboration is proving to be a game-changer in developing scalable solutions to support long-term preservation and ensure collections remain reliably accessible for future generations.

The web archiving process is not a one-off action. A suite of applications is typically deployed to support different stages of the process, though they may be integrated into a single end-to-end workflow. Much of the software is available as open source, allowing institutions free access to the source code for use and/or modification at no cost.

Integrated Systems for Web-archiving

A small number of integrated systems are available for those with sufficient technical staff to install, maintain and administer a system in-house. These typically offer integrated web archiving functionality across most of the life cycle, from selection and permissions management to crawling, quality assurance, and access. Three are featured here.

PANDAS

PANDAS (PANDORA Digital Archiving System) was one of the first available integrated web archiving systems. First implemented by the National Library of Australia (NLA) in 2001, PANDAS is a web application written in Java and Perl that provides a user-friendly interface to manage the web archiving workflow. It supports selection, permissions, scheduling, harvests, quality assurance, archiving, and access. PANDAS is not open source software, though it has been used by other institutions (most notably the UK Web Archiving Consortium from 2004 to 2008). It is used by the NLA for selective web archiving, whilst the Internet Archive supports their annual snapshots of the Australian domain.

Web Curator Tool (WCT)

The Web Curator Tool is an open source workflow tool for managing the selective web archiving process, developed collaboratively by the National Library of New Zealand and the British Library with Oakleigh Consulting. It supports selection, permissions, description, harvests, and quality assurance, with a separate access interface. WCT is written in Java within a flexible architecture and is publicly available for download from SourceForge under an Apache public licence. The WCT website is the hub for the developer

community and there are active mailing lists for both users and developers. The highly modular nature of the system minimizes system dependencies.

NetarchiveSuite

NetarchiveSuite is a web archiving application written in Java for managing selective and broad domain web archiving, originally developed in 2004 by the two legal deposit libraries in Denmark (Det Kongelige Bibliotek and Statsbiblioteket). It became open source in 2007 and has received additional development input from the Bibliothèque nationale de France and the Österreichische Nationalbibliothek since 2008. It is freely available under the GNU Lesser General Public License (LGPL). The highly modular nature of the system enables flexible implementation solutions.

Third party and commercial services

Third party commercial web archiving services are increasingly used by organizations that prefer not to establish and maintain their own web archiving technical infrastructure. The reasons behind this can vary widely. Often it is not simply about the scale of the operation or the perceived complexity, but the business need and focus. Many organizations do not wish to invest in any skills or capital that is not core to their business. Others may use such a service to avoid capital investment. Moreover, organizations are increasingly moving their computing and IT operations into the cloud, or using a SAAS (Software as a Service) provider. Web archiving is no exception. From a legal and compliance perspective, third party services are sometimes preferred as they can provide not just the technology but also the skills and support required to meet business needs. This section introduces some of the third party services currently available but is of course a non-exhaustive list, and inclusion here should not be taken as recommendation.

Archive-It

Archive-It is a subscription web archiving service provided by the Internet Archive. Customers use the service to establish specific collections, for example about the London 2012 Olympics, government websites, human rights, and course reading lists. A dedicated user interface is provided for customers to select and manage seeds, set the scope of a crawl and crawl frequency, monitor crawl progress and perform quality assurance, add metadata and create landing pages for their collections. Collections are made public by default via the Archive-It website, with private collections requiring special arrangement. The access interface supports both URL and full text searching. Over 200 partners use the service, mostly from the academic or cultural heritage sectors. The cost of the service depends on the requirements of the collecting institution

Archivethe.Net

Archivethe.Net is a web-based web archiving service provided by the Internet Memory Foundation (IMF). It enables customers to manage the entire workflow via a web interface to three main modules: Administration (managing users), Collection (seed and crawl management), and Report (reports and metrics at different levels). The platform is available in both English and French. Alongside full text searching and collection of multimedia content, it also supports an automated redirection service for live sites. Automated QA tools are being developed though IMF can also provide manual quality assurance services, as well as direct collection management for institutions not wishing to use the online tool. Costs are dependent upon the requirements of the collecting institution. Collections can be made private or remain openly accessible,

in which case they may be branded as required by the collecting institutions and appear in the IMF collection. The hosting fee in such cases is absorbed by IMF.

The University of California's Curation Centre (UC3)

As part of the California Digital Library, provides a fully hosted Web Archiving Service for selective web archive collections. University of California departments and organizations are charged only for storage. Fees are levied for other groups and consortia, comprising an annual service fee plus storage costs. Collections may be made publicly available or kept private. Around 20 partner organizations have made collections available to date. Full text search is provided and presentation of the collections can be branded as required by collecting institutions.

Private companies

Private companies offer web archiving services particularly tailored to business needs. Hanzo Archives, for example, provide a commercial website archiving service to meet commercial business needs around regulatory compliance, e-discovery and records management. Hanzo Archives emphasize their ability to collect rich media sites and content that may be difficult for a standard crawler to pick up, including dynamic content from Sharepoint, and wikis from private internets, alongside public and private social media channels. (More details about the possibilities afforded by the Hanzo Archives service can be found in the Coca-Cola case study) Similarly, Reed Archives provide a commercial web archiving service for organizational regulatory compliance, litigation protection, eDiscovery and records management. This includes an 'archive-on-demand' toolset for use when browsing the web. In each case, the cost of the service is tailored to the precise requirements of the customer. Other companies and services are also available and readers are encouraged to search online for further options should such a service be of interest.

Case study 1: The UK Web Archive

The UK Web Archive (UKWA) was established in 2004 by the UK Web Archiving Consortium. It was originally a six-way partnership, led by the British Library in conjunction with the Wellcome Library, Jisc, the National Library of Wales, the National Library of Scotland and The National Archives (UK).

UKWA partners select and nominate websites using the features of the web archiving system hosted on the UK Web Archive infrastructure maintained by the British Library. The British Library works closely with a number of other institutions and individuals to select and nominate websites of interest. Selectively archived websites are revisited at regular intervals so that changes over time are captured.

The technical infrastructure underpinning the UK Web Archive is managed by the British Library. The Archive was originally established with the PANDAS software provided by the National Library of Australia, hosted by an external agency, but in 2008 the archive was moved in-house and migrated into the Web Curator Tool (WCT) system.

A customized version of the Wayback interface developed by the Internet Archive is used as the WCT front end and provides searchable access to all publicly available archived websites. Full text searching is enabled in addition to standard title and URL searches and a subject classification schema. The web archiving team at the library have recently released a number of visualization tools to aid researchers in understanding and finding content in the collection.

Special collections have been established on a broad range of topics. Many are subject based, for example the mental health and the Free Church collections. Others document the online response to a notable event in recent history, such as the UK General Elections, Queen Elizabeth II's Diamond Jubilee and the London 2012 Olympics.

Many more single sites, not associated with a given special collection, have been archived on the recommendation of subject specialists or members of the public. These are often no longer available on the live web, for example the website of UK Member of Parliament Robin Cook or Antony Gormley's One & Other public art project, acquired from Sky Arts.

Case study 2: The Internet Memory Foundation

The Internet Memory Foundation (IMF) was established in 2004 as a non-profit organization to support web archiving initiatives and develop support for web preservation in Europe. Originally known as the European Archive Foundation, it changed its name in 2010. IMF provides customers with an outsourced fully fledged web archiving solution to manage the web archiving workflow without them having to deal with operational workflow issues.

IMF collaborates closely with Internet Memory Research (IMR) to operate a part of its technical workflows for web archiving. IMR was established in 2011 as a spin off from the IMF. Both IMF and IMR are involved in research projects that support the growth and use of web archives.

IMR provides a customizable web archiving service, Archivethe.Net (AtN). AtN is a shared web-archiving platform with a web-based interface that helps institutions to easily and quickly start collecting websites including dynamic content and rich media. It can be tailored to the needs of clients, and institutions retain full control of their collection policy (ability to select sites, specify depth, gathering frequency, etc.). Quality control services can be provided on request. Most is done manually in order to meet high levels of institutional quality requirements, and IM has a dedicated QA team composed of QA assessors. IM has developed a methodology for visual comparison based on tools used for crawling and accessing data, though they are also working on improving tools and methods to deliver a higher initial crawl quality.

Partner institutions, with openly accessible collections for which the IM provides a web archiving service, include the UK National Archives and the UK Parliament.

Access to publicly available collections is provided via the IM website. IM provides a full text search facility for most of its online collections, in addition to URL-based search. Full text search results can be integrated on a third party website and collections can be branded by owners as necessary.

Following the architecture of the Web Continuity Service by The National Archives ([The National Archives, 2010](#)), IM implemented an 'automatic redirection service' to integrate web archives with the live web user experience. When navigating on the web, users are automatically redirected to the web archive if the resource requested is no longer available online. Within the web archive, the user is pointed to the most recent crawled instance of the requested resource. Once the resource is accessed, any link on the page will send the user back to the live version of the site. This service is considered to increase the life of a link, to improve users' experience, online visibility and ranking, and to reduce bounce rates.

Web archiving collections are available for public browsing from the IM website, a combination of both domain and selective collections from its own and from partner institutions.

Case study 3: The Coca-Cola web archive

The Coca-Cola Web Archive was established to capture and preserve corporate Coca-Cola websites and social media. It is part of the Coca-Cola Archive, which contains millions of both physical and digital artefacts, from papers and photographs to adverts, bottles, and promotional goods. Coca-Cola's online presence is vast, including not only several national Coca-Cola websites but also for example, the Coca-Cola Facebook page and Twitter stream, and other Coca-Cola owned brands (500 in all). The first Coca-Cola website was published in 1995.

Since 2009, Coca-Cola has collaborated with Hanzo Archives and now utilizes their commercial web archiving service. Alongside the heritage benefits of the web archive, the service also provides litigation support where part or all of the website may be called upon as evidence in court and regulatory compliance for records management applications.

The Coca-Cola web archive is a special themed web archive that contains all corporate Coca-Cola sites and other specially selected sites associated with Coca-Cola. It is intended to be as comprehensive as possible, with integrity/functionality of captured sites of prime importance. This includes social media and video, whether live-streamed or embedded (including Flash). Artefacts are preserved in their original form wherever possible, a fundamental principle for all objects in the Coca-Cola Archive.

Hanzo Archives' crawls take place quarterly and are supplemented by occasional event-based collection crawls, such as the 125th anniversary of Coca-Cola, celebrated in 2011. Hanzo's web archiving solution is a custom-built application. Web content is collected in its native format by the Hanzo Archives web crawler, which is deployed to the scale necessary for the task in hand.

Quality assurance is carried out with a two-hop systematic sample check of crawl contents that forces use of the upper-level navigation options and focuses on the technical shape of the site.

The Archive is currently accessible only to Coca-Cola employees, on a limited number of machines. Remote access is provided by Hanzo using their own access interface. Proxy-based access ensures that all content is served directly from the archive and that no 'live-site leakage' is encountered. The archive may be made publicly accessible in the future inside The World of Coca-Cola, in Atlanta, Georgia, USA.

The Coca-Cola web archive collection contains over six million webpages and over 2TB of data. Prior to their collaboration with Hanzo, early attempts at archiving resulted in incomplete captures so early sites are not as complete as the company would like. The collection also contains information about many national and international events for which Coca-Cola was a sponsor, including the London 2012 Olympics and Queen Elizabeth II's Diamond Jubilee.

Conclusions

Web archiving technology has significantly matured over the past decade, as has our understanding of the issues involved. Consequently we have a broad set of tools and services which enable us to archive and preserve aspects of our online cultural memory and comply with regulatory requirements for capturing and preserving online records. The work is ongoing, for as long as the Internet continues to evolve, web archiving technology must evolve to keep pace.

Alongside technical developments, the knowledge and experience gained through practical deployment and use of web archiving tools has led to a much better understanding of best practices in web archiving, operational strategies for embedding web archiving in an organizational context, business needs and benefits, use cases, and resourcing options. Organizations wishing to embark on a web archiving initiative must be very clear about their business needs before doing so. Business needs

should be the fundamental driver behind any web archiving initiative and will significantly influence the detail of a resulting web archiving strategy and selection policy. The fact that commercial services and technologies have emerged is a sign of the maturity of web archiving as a business need, as well as a discipline.

Resources



Pennock, M., 2013. Web-Archiving, DPC Technology Watch Report 13-01 March 2013

<http://dx.doi.org/10.7207/twr13-01>

This report is intended for those with an interest in, or responsibility for, setting up a web archive. It introduces and discusses the key issues faced by organizations engaged in web archiving initiatives, whether they are contracting out to a third party service provider or managing the process in-house and provides a detailed overview of the main software applications and tools currently available.

ISO, 2012, ISO 28500:2009 Information and Documentation – the WARC file format

http://www.iso.org/iso/catalogue_detail.htm?csnumber=44717

The WARC (Web ARChive) format is a container format for archived websites, also known as ISO 28500:2009. It is a revision of the Internet Archive's ARC File Format used to store web crawls harvested from the World Wide Web.

ISO, 2013 ISO/TR 14873:2013 Information and Documentation – Statistics and quality issues for web archiving

http://www.iso.org/iso/catalogue_detail.htm?csnumber=55211

This technical report defines statistics, terms and quality criteria for Web archiving. It considers the needs and practices across a wide range of organisations such as libraries, archives, museums, research centres and heritage foundations.

Meyer E 2010 (a), Researcher Engagement with Web Archives: State of the Art Report, JISC

<http://ie-repository.jisc.ac.uk/544/>

This report summarizes the state of the art of web archiving in relationship to researchers and research needs focussing primarily on individual researchers and institutions.

Perma: Scoping and Addressing the Problem of Link and Reference Rot in Legal Citations

Zittrain, Jonathan and Albert, Kendra and Lessig, Lawrence, Perma: Scoping and Addressing the Problem of Link and Reference Rot in Legal Citations (October 1, 2013). Harvard Public Law Working Paper No. 13-42. Available at SSRN: <http://ssrn.com/abstract=2329161> or <http://dx.doi.org/10.2139/ssrn.2329161>

http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2329161 or <http://dx.doi.org/10.2139/ssrn.2329161>

This article from the Perma project team documents a serious problem of reference rot: more than 70% of the URLs within the Harvard Law Review and other journals, and 50% of the URLs found within United States Supreme Court opinions, do not link to the originally cited information. It proposes a solution for authors and editors of new scholarship that involves libraries undertaking the distributed, long-term preservation of link contents.

Scholarly Context Not Found: One in Five Articles Suffers from Reference Rot

<http://dx.doi.org/10.1371/journal.pone.0115253>

This large-scale study looked into approximately 600K links extracted from over 3M scholarly papers published between 1997 and 2012. Those were links to so-called web-at-large resources, i.e. not links to other scholarly papers. It found one out of five STM articles suffering from reference rot, meaning it is impossible to revisit the web context that surrounds them some time after their publication. When only considering STM articles that contain references to web resources, this fraction increases to seven out of ten.

The National Archives, 2010. Government Web Archive: Redirection Technical Guidance for Government Departments, version 4.2, The National Archives (UK)

<http://www.nationalarchives.gov.uk/documents/information-management/redirection-technical-guidance-for-departments-v4.2-web-version.pdf>

This guidance describes an innovative service that provides URL rewriting and redirection functionality for UK Government web pages by setting up redirection to the UK Government web archive where a requested URL does no longer exists on a departmental web site.



MEMENTO and the Time Travel Service

<http://www.mementoweb.org/>

Memento is a tool which allows users to see a version of a web resource as it existed at a certain point in the past. It is now used in several web archives. The Time Travel service based on Memento checks a range of servers including many web archives and tries to find a web page as it existed around the time of your choice.

Archive-It

<http://www.archive-it.org/>

Archivethe.Net

<http://www.archivethe.net/en/>

Hanzo Archives

<http://www.hanzoarchives.com/>

Internet Memory Foundation & Internet Memory Research

<http://www.internetmemory.org/en/>

Wayback

<http://www.sourceforge.net/projects/archive-access/files/wayback/>

Netarchive Suite

<https://sbforge.org/display/NAS/NetarchiveSuite>

PANDAS

<http://pandora.nla.gov.au/pandas.html>

Reed Archives

<http://www.reedarchives.com/>

UC3 Web Archiving Service

<http://www.cdlib.org/services/uc3/was.html>

Web Curator Tool

<http://webcurator.sourceforge.net/>



International Internet Preservation Consortium

<http://www.netpreserve.org>

The IIPC is a membership organization dedicated to improving the tools, standards and best practices of web archiving while promoting international collaboration and the broad access and use of web archives for research and cultural heritage. There are many valuable resources on the website including excellent short videos such as the example below.



Why Archive the Web?

<https://www.youtube.com/watch?v=pU32rjTaMFE>

A short video published on 18 Oct 2012 introducing the challenges of web-archiving and the IIPC. (2 mins 53 secs).

What is a Web Archive?

<https://youtu.be/ubDHY-ynWi0>

This short video explains 'Web Archiving' and why it is important that the UK Legal Deposit libraries support it. It was produced as part of the Arts and Humanities Research Council funded 'Big UK Domain Data for the Arts and Humanities' project. (2 mins 31 secs)

What do the UK Web Archive collect?

<https://youtu.be/1QLMPiRwJEo>

This video for users explains what they can expect to find and where they might go to access the three collections that the UK Web Archive hold. It was produced as part of the Arts and Humanities Research Council funded 'Big UK Domain Data for the Arts and Humanities' project. (2 mins 55 secs)

Further case studies



NDSA Website content case studies

The US National Digital Stewardship Alliance (NDSA) examines the value, opportunities and obstacles for selective preservation of the following specific web content types:

Science, Medicine, Mathematics, and Technology forums

http://www.digitalpreservation.gov/ndsa/working_groups/documents/ScienceForums_CaseStudy_public_v2.pdf

December 2013 (3 pages).

Science, Medicine, Mathematics, and Technology blogs

http://www.digitalpreservation.gov/ndsa/working_groups/documents/ScienceBlogs_CaseStudy_public_v2.pdf

December 2013 (3 pages).

Born-Digital Community and Hyperlocal News

http://www.digitalpreservation.gov/ndsa/working_groups/documents/NDSA_CaseStudy_CommunityNews.pdf

February 2013 (3 pages).

Citizen Journalism

http://www.digitalpreservation.gov/ndsa/working_groups/documents/NDSA_CaseStudy_CitizenJournalism.pdf

February 2013 (3 pages).

On the Development of the University of Michigan Web Archives: Archival Principles and Strategies

<http://files.archivists.org/pubs/CampusCaseStudies/Case13Final.pdf>

Michael Shallcross, Bentley Historical Library, University of Michigan details the strategies and procedures the University Archives and Records Program (UARP) followed to develop its collection of archived websites, and how it initiated a large-scale website preservation project as part of a broader effort to proactively capture and maintain select electronic records of the University. 2011 (29 pages).

References

Pennock, M., 2013. Web-Archiving, *DPC Technology Watch Report 13-01* March 2013. Available: <http://dx.doi.org/10.7207/twr13-01>

The National Archives, 2010. *Government Web Archive: Redirection Technical Guidance for Government Departments*, version 4.2, The National Archives (UK). Available: <http://www.nationalarchives.gov.uk/documents/information-management/redirection-technical-guidance-for-departments-v4.2-web-version.pdf>

Digital Preservation **Handbook**

Technical Solutions and Tools



Illustration by Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Who is it for?

Operational managers (DigCurV Manager Lens) and staff (DigCurV Practitioner Lens) in repositories, publishers and other data creators, third party service providers.

Assumed level of knowledge

Novice to Intermediate.

Purpose

- To focus on technical tools and applications that support digital preservation: software, applications, programs and technical services.
- To consider the practical deployment of preservation techniques and technologies whether as relatively small and discrete programs (like DROID) or enterprise wide solutions that integrate many tools.
- This section excludes other more strategic or policy issues and standards that are sometimes described as tools: these are covered elsewhere in the Handbook.

Gold sponsor



Silver sponsors



Bronze sponsors



Reusing this information

You may re-use this material in English (not including logos) with required acknowledgements free of charge in any format or medium. See [How to use the Handbook](#) for full details of licences and acknowledgements for re-use.

For permission for translation into other languages email: handbook@dpconline.org

Please use this form of citation for the Handbook: Digital Preservation Handbook, 2nd Edition, <http://handbook.dpconline.org/>, Digital Preservation Coalition © 2015.

Contents

Tools	4
Resources	6
Case studies.....	8
Fixity and checksums.....	9
Resources	12
References.....	13
File formats and standards.....	13
Resources	18
Case studies.....	20
Information security.....	21
Resources	23
Case studies.....	25
References.....	25
Cloud services.....	26
Resources	29
Case studies.....	30
Digital forensics	32
Resources	34
Case studies.....	35
References.....	36
Persistent identifiers	36
Resources	38
Case studies.....	40

Tools



Illustration by Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

A beginner's guide to digital preservation tools

The utility of technical tools for digital preservation depends on the context of their deployment. A community recommendation may be strong but if it does not align with your specific function or organisational context then there is a significant chance that the tool will fail to perform. So before selecting digital preservation tools it is important to consider carefully the technical workflow and institutional setting in which they are embedded. A practical example of this has been presented by Northumberland Estates who developed a straightforward evaluation framework to assess tools in context.

An alternative way to consider this topic is to review the extent to which any given tool will deliver preservation actions arising from an agreed preservation plan, which in turn derives from a given policy framework.

Thinking about digital preservation tools

The following issues are frequently encountered in the process of deploying digital preservation tools. This is not a comprehensive list but consideration of these issues will help sensible and realistic choices.

Open source versus commercial software

Some organizations - often in higher education and especially institutional research repositories - are comfortable with the use of open source software, especially where they have an in-house group of developers. 'Open source' software is where the underlying code is made available for free, enabling a free flow of additions, amendments or development. Other organizations which don't have easy access to developers, tend to have procurement rules that prefer 'off-the shelf' commercial solutions backed by on-going support contracts. The distinction between Open Source versus Commercial software is often over-stated because both influence each other. Nonetheless you may need to consider your organization's norms and culture while you select tools.

Enterprise-level solutions versus micro-services

Some digital preservation tools are designed to offer 'soup to nuts' solutions, meaning that they provide an integrated end-to-end process that enables all (or most) digital preservation functions to

be delivered for a whole organisation. In fact enterprise-level solutions are most often constructed by aggregating individual tools integrated into a single interface. The solution to any given problem might be relatively simple and your organisation may be happy assembling a series of small tools for discrete functions. This encourages rapid progress and is helpful with testing and trialling tools; but it can be hard to maintain over an extended period. In other organisations there is much tighter control over the deployment of software and an expectation that solutions are built across an entire workflow - requiring comprehensive solutions. This can be slower to respond but can be more sustainable in the long term. Before selecting a tool it is helpful to consider where on this spectrum your organization normally sits.

Describing workflows

A key consideration for tools is where they sit on an overall workflow so before selecting tools it helps to consider and map out the entire workflow. Being explicit about a workflow can also help identify redundant processes as well major bottlenecks. One frequent challenge is that tools solve a problem in one element of a workflow, only to create a problem elsewhere. In addition, organisations may have multiple workflows that may have different requirements that conflict in some way. Describing a workflow therefore provides a basis for anticipating difficulties and can provide a roadmap for ongoing development.

Specifying clear requirements

In order to evaluate the usefulness and value to your organisation of the many tools available it helps to have an explicit statement of requirements. Tools can be compared and benchmarked transparently and decisions justified accordingly. Properly executed, requirements-gathering activities can involve a range of stakeholders and therefore maximise the potential for alignment and efficiency, achieving wider strategic and organisational objectives.

Changing and evolving requirements

It is normal for requirements to change through time. Indeed digital preservation is largely concerned with meeting the challenges associated with inevitable changes in technology. So it is necessary to monitor and review tools to ensure that they remain fit for purpose and that any changes in requirements are made explicit. A periodic review of the specification of requirements is recommended.

Sustainability of tools and community participation

An important consideration in any decision over the tools you use for digital preservation is the sustainability element. Sustainability in terms of tools may include an active user base, support, and development. For instance, a large user base, both in terms of commercial and open source providers can be a vital indicator for identifying a viable tool. It's worth noting that a community can change rapidly and for reasons that might not be easily predicted. 'New kids on the block' can quickly become mainstream while large communities can dwindle as quickly as new technologies overtake existing ones. Consequently it may be necessary to monitor the health of the developer community supporting your tools.

Finding digital preservation tools: tools and tools registries

One of the welcome features of digital preservation in the last two decades has been the rapid development of software, tools and services that enhance and enable digital preservation workflows. As the digital preservation community has grown in size and sophistication so our tools have become more powerful and more refined. This proliferation and increased specialism can also act as a barrier to deployment: especially when tools have been the product of relatively short lived research projects

with limited reach. Consequently the diversity of tools can seem increasingly bewildering to new users, while the route to market for developers is increasingly complicated.

Tools registries have emerged in recent years as a way to help users find tools that they need. A number of registries now exist that describe digital preservation tools. Depending on the interests of the people behind them, they can also provide detailed descriptions, reviews or comments about tools from the wider community. So they are not just helpful for users: by allowing experts to review tools and assess their performance they signpost strengths and weaknesses and provide a basis for future development; by connecting tools to users they help developers reach a much wider audience and get feedback to improve their tools.

Registries are a common way for the digital preservation community to share information. Other types of registries exist such as 'format registries' that outline the performance of given file formats, or 'environment registries' that describe the technology stack necessary to create an execution environment to emulate or virtualize software. These are covered elsewhere in the Handbook.

Too many registries?

While registries are a good way to manage the proliferation of tools, it is now recognised that a proliferation of registries is also a potential barrier to use. The [COPTR](#) registry was designed specifically to address this problem, drawing on data from multiple sources including DCC, POWRR, and the Library of Congress.

Practical support and guidance

Having considered some of the tools registries and digital preservation tools that are available to organisations, the next question that often arises is which one to choose that fits your organisational purpose. First and foremost it is important that your selection is aligned to organisational need and strategic direction; the resources and case studies below provide evaluation tools and advice to support successful implementation.

Resources



Tool registries

Community Owned digital Preservation Tool Registry COPTR

http://coptr.digipres.org/Main_Page

COPTR describes tools useful for long term digital preservation and acts primarily as a finding and evaluation tool to help practitioners find the tools they need to preserve digital data. COPTR aims to collate the knowledge of the digital preservation community on preservation tools in one place. It was initially populated with data from registries run by the COPTR partner organisations, including those maintained by the Digital Curation Centre, the Digital Curation Exchange, National Digital Stewardship Alliance, the Open Preservation Foundation, Preserving digital Objects With Restricted Resources project (POWRR) <http://digitalpowrr.niu.edu/> listed below. COPTR captures basic, factual details about a tool, what it does, how to find more information (relevant URLs) and references to user experiences with the tool. The scope is a broad interpretation of the term "digital preservation". In other words, if

a tool is useful in performing a digital preservation function such as those described in the OAIS model or the DCC lifecycle model, then it's within scope of this registry.

APARSEN tools registry

<http://www.alliancepermanentaccess.org/index.php/tools/tools-for-preservation/>

The APARSEN tools repository attempts to build an evidence-base for preservation tools, and in particular to try to identify which tools are appropriate for which type of data. APARSEN collects details of preservation related software, examples of data, and the evidence of preservation linking software to types of data. Some of this evidence comes from specific testbeds but much comes from user scenarios. The resource is now maintained by the Alliance for Permanent Access (APA).

AV Preserve tools list

<http://www.avpreserve.com/avpsresources/tools/>

A list of tools of particular use in the long term preservation of audio visual materials, both digitised and born-digital.

Digital Curation Centre (DCC) tools and services list

<http://www.dcc.ac.uk/resources/external/tools-services>

The DCC is a centre of excellence, to support researchers in the UK tackling challenges for the preservation and curation of digital resources. To achieve this goal it offered a number of support and advisory services supported with targeted research and development. The former includes a catalogue of tools and services which categorises tools for researchers and curators. The information is also integrated in COPTR (see above).

DCH-RP registry

<http://www.dch-rp.eu/index.php?en/137/registry-of-services-tools>

The Digital Cultural Heritage Roadmap for Preservation (DCH-RP) tools registry collected and described information and knowledge related to tools, technologies and systems that can be applied for the purposes of digital cultural heritage preservation. Version 3 of the registry was created in 2014.

Inventory of FLOSS (Free/libre open-source software) in the cultural heritage domain

https://docs.google.com/spreadsheet/ccc?key=0Ag_7rVJwt0CpdFRJOEJxdEk4ZEMxQ01jaDgxQXFSTkE#gid=0

Produced by the EU funded Europeana Project, this inventory lists free open source software which may be of use in the cultural heritage sector. While not limited to digital preservation tools the inventory does contain information on a variety of tools with digital preservation applications, assessing their purpose, quality of documentation, level of support, license requirements and providing links to project information and source code. Background information on FLOSS is available on the Europeana site <http://www.europeana.eu/portal/>.

Library of Congress NDIIPP tools showcase

<http://www.digitalpreservation.gov/tools/>

The Library of Congress's digital preservation tools registry is a selective list of tools and services of interest to those working in digital preservation. It is no longer being actively maintained and content is integrated in COPTR (see above).

Preserving digital Objects With Restricted Resources (POWRR) Tool Grid

<http://digitalpowrr.niu.edu/tool-grid/>

POWRR investigated, evaluated, and recommended scalable, sustainable digital preservation solutions for organisations with relatively small amounts of data and/or fewer resources. A significant output of the project was the tool grid produced in early 2013 based on the OAIS Reference Model functional categories. An up to date version of the POWRR Tool Grid can now be generated in COPTR (see above).



Digital Preservation Q&A

<http://qanda.digipres.org/>

This is a site where you can post queries and answers to help each other make best use of tools, techniques, processes, workflows, practices and approaches to insuring long term access to digital information. Digital Preservation Q&A is currently moderated by representatives from NDSA and OPF member organizations.

Practical e-records

<http://e-records.chrisprom.com/author/prom/>

Software and Tools for Archivists blog from Chris Prom. Although some information may be several years old the blog provides a useful starting point for understanding the uses of a variety of tools for digital preservation and a standardised evaluation of the tools against set criteria, including ease of installation, usability, scalability etc. In addition to information on tools the blog contains a host of other useful resources, including policy and workflow templates, recommended approaches.

Case studies



Diary of a repository preservation project

<http://blog.soton.ac.uk/keepit/>

A record of progress (between April 2009 and September 2010) as the Jisc-funded KeepIt project tackled the challenges of preserving digital repository content in research, teaching, science and the arts. It includes helpful experience for assessing preservation tools.

Northumberland Estates

http://wiki.dpconline.org/index.php?title=Northumberland_estates_case_study

Northumberland Estates developed a straightforward evaluation framework to assess tools in context. The project set out to survey digital repository options currently available for small to medium organisations with limited resources. Note the recommendations reached in the final business case reflect the organisational needs of Northumberland Estates and may not align themselves with your own goals. The case study was prepared as part of the Jisc-funded SPRUCE project.

Fixity and checksums



Illustration by Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Fixity

“Fixity, in the preservation sense, means the assurance that a digital file has remained unchanged, i.e. fixed.” ([Bailey, 2014](#)). Fixity doesn’t just apply to files, but to any digital object that has a series of bits inside it where that ‘bitstream’ needs to be kept intact with the knowledge that it hasn’t changed. Fixity could be applied to images or video inside an audiovisual object, to individual files within a zip, to metadata inside an XML structure, to records in a database, or to objects in an object store. However, files are currently the most common way of storing digital materials and fixity of files can be established and monitored through the use of checksums.

Checksums

A checksum on a file is a ‘digital fingerprint’ whereby even the smallest change to the file will cause the checksum to change completely. Checksums are typically created using cryptographic techniques and can be generated using a range of readily available and open source tools. It is important to note that whilst checksums can be used to detect if the contents of a file have changed, they do not tell you where in the file that the change has occurred.

Checksums have three main uses:

1. To know that a file has been correctly received from a content owner or source and then transferred successfully to preservation storage
2. To know that file fixity has been maintained when that file is being stored.
3. To be given to users of the file in the future so they know that the file has been correctly retrieved from storage and delivered to them.

This allows a 'chain of custody' to be established between those who produce or supply the digital materials, those responsible for its ongoing storage, and those who need to use the digital material that has been stored. In the OAIS reference model ([ISO, 2012](#)) these are the producers, the OAIS itself is the repository, and the consumers.

Application in digital preservation

If an organisation has multiple copies of their files, for example as recommended in the [Storage](#) section, then checksums can be used to monitor the fixity of each copy of a file and if one of the copies has changed then one of the other copies can be used to create a known good replacement. The approach is to compute a new checksum for each copy of a file on a regular basis and compare this with the reference value that is known to be correct. If a deviation is found then the file is known to have been corrupted in some way and will need replacing with a new good copy. This process is known as 'data scrubbing'.

Checksums are ideal for detecting if unwanted changes to digital materials have taken place. However, sometimes the digital materials will be changed deliberately, for example if a file format is migrated. This causes the checksum to change. This requires new checksums to be established after the migration which become the way of checking data integrity of the new file going forward.

Files should be checked against their checksums on a regular basis. How often to perform checks depends on many factors including the type of storage, how well it is maintained, and how often it is being used. As a general guideline, checking data tapes might be done annually and checking hard drive based systems might be done every six months. More frequent checks allow problems to be detected and fixed sooner, but at the expense of more load on the storage system and more processing resources.

Checksums can be stored in a variety of ways, for example within a [PREMIS](#) record, in a database, or within a 'manifest' that accompanies the files in a storage system.

Tool support is good for checksum generation and use. As they are relatively simple functions, checksums are integrated into many other digital preservation tools. For example, generating checksums as part of the ingest process and adding this fixity information to the Archive Information Packages generated, or allowing manifests of checksums to be generated for multiple files and for the manifest and files to be bundled together for easy transport or storage. In addition md5sum and md5deep provide simple command line tools that operate across platforms to generate checksums on individual files or directories.

There are several different checksum algorithms, e.g. MD5 and SHA-256 that can be used to generate checksums of increasing strength. The 'stronger' the algorithm then the harder it is to deliberately change a file in a way that goes undetected. This can be important for applications where there is a need to demonstrate resistance to malicious corruption or alteration of digital materials, for example where evidential weight and legal admissibility is important. However, if checksums are being used to detect accidental loss or damage to files, for example due to a storage failure, then MD5 is sufficient and has the advantage of being well supported in tools and is quick to calculate.

The Handbook follows the National Digital Stewardship Alliance (NDSA) preservation levels ([NDSA, 2013](#)) in recommending four levels at which digital preservation can be supported through file fixity and data integrity techniques. Many of the benefits of fixity checking can only be achieved if there are multiple copies of the digital materials, for example allowing repair if integrity of one of the copies has been lost.

Level	Activity	Risks addressed and benefits achieved
1	<ul style="list-style-type: none"> • Check file fixity on ingest if it has been provided with the content. • Create fixity info if it wasn't provided with the content. 	<ul style="list-style-type: none"> • Corrupted or incorrect digital materials are not knowingly stored. • Authenticity of the digital materials can be asserted. • Baseline fixity established so unwanted data changes have potential to be detected.
2	<ul style="list-style-type: none"> • Check fixity on all ingests • Use write-blockers when working with original media • Virus-check high risk content. 	<ul style="list-style-type: none"> • No digital material of unconfirmed integrity can enter preservation storage. Evidential weight supported for authenticity. • Assurance can be given to all content providers that their content has been safely received. Original media is protected. • No malicious content can enter preservation storage.
3	<ul style="list-style-type: none"> • Check fixity of content held on preservation storage systems at regular intervals. • Maintain logs of fixity info and supply audit on demand. • Ability to detect corrupt data. • Virus-check all content. 	<ul style="list-style-type: none"> • Protection from wide range of data corruption and loss events. Problems with storage are detected earlier. • Data corruption or loss does not go undetected due to 'silent errors' or 'undetected failures'. Digital materials are not in a state of 'unknown' integrity. • Ongoing evidential weight can be given that digital materials are intact and correct.
4	<ul style="list-style-type: none"> • Check fixity of all content in response to specific events or activities • Ability to replace/repair corrupted data • Ensure no one person has write access to all copies. 	<ul style="list-style-type: none"> • Failure modes that threaten digital materials are proactively countered. All copies of digital materials are actively maintained. • Assurance to users of the integrity and authenticity of digital materials being accessed. • Effectiveness of preservation approach can be measured and demonstrated. • Compliance with standards, e.g. ISO 16363 Audit and certification of trustworthy digital repositories.

Write-blocking

Note that the National Digital Stewardship Alliance (NDSA) recommends the use of write-blockers at level 2. This is to prevent write access to media that digital materials might be on prior to being copied

to the preservation storage system. For example, if digital material is delivered to an organisation on a hard disc drive or USB key then a write blocker would prevent accidental deletion of this digital material when the drive or key is read. Digital material might not be on physical media, e.g. it could be on a legacy storage server or delivered through a network transfer, e.g. an ftp upload. In these cases write blockers wouldn't apply and other measures would be used to make the digital material 'read only' on the source and hence immutable before confirmation that the digital material has been successfully transferred to preservation storage. Write blockers also don't exist for all types of media. If a write-blocker is applicable then the costs/skills required to use them should be balanced against the risk of damage to the original digital material or the need to have rigorous data authenticity. Therefore, some organisations might consider use of write blockers to be unnecessary or a level 3 or level 4 step.

Resources



Bailey, J., 2014, Protect Your Data: File Fixity and Data Integrity, The Signal, Library of Congress.

<http://blogs.loc.gov/digitalpreservation/2014/04/protect-your-data-file-fixity-and-data-integrity/>

Checking Your Digital Content: What is Fixity and When Should I Be Checking It?

http://digitalpreservation.gov/ndsa/working_groups/documents/NDSA-Fixity-Guidance-Report-final100214.pdf?loclr=blogsig

Many in the preservation community know they should be checking the fixity of their content, but how, when and how often? This document published by NDSA in 2014 aims to help stewards answer these questions in a way that makes sense for their organization based on their needs and resources (7 pages).



AVPreserve Fixity Tool

<http://www.avpreserve.com/tools/fixity/>

MD5

<https://tools.ietf.org/html/rfc1321>

SHA-1

<http://csrc.nist.gov/publications/fips/fips180-4/fips-180-4.pdf>

SHA-256

<http://csrc.nist.gov/groups/STM/cavp/documents/shs/sha256-384-512.pdf>

Md5deep and hashdeep

http://coptr.digipres.org/Md5deep_and_hashdeep

md5sum

http://coptr.digipres.org/Md5sum_Unix_command



The "Checksum" and the Digital Preservation of Oral History

https://www.youtube.com/watch?v=Emom_ncMqu0

A good short overview not limited to oral history, this video provides a brief introduction to the role of the checksum in digital preservation. It features Doug Boyd, Director of the Louie B. Nunn Center for Oral History at the University of Kentucky Libraries. (3 mins 25 secs)

References

Bailey, J., 2014. Protect Your Data: File Fixity and Data Integrity. *The Signal*. [blog]. Available:

<http://blogs.loc.gov/digitalpreservation/2014/04/protect-your-data-file-fixity-and-data-integrity/>

ISO, 2012. ISO 14721:2012 - *Space Data and Information Transfer Systems – Open Archival Information System (OAIS) – Reference Model, 2nd edn*. Geneva: International Organization for Standardization.

Available:

http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=57284

NDSA , 2013. *The NDSA Levels of Digital Preservation: An Explanation and Uses, version 1 2013*.

National Digital Stewardship Alliance. Available:

http://www.digitalpreservation.gov/ndsa/working_groups/documents/NDSA_Levels_Archiving_2013.pdf

File formats and standards

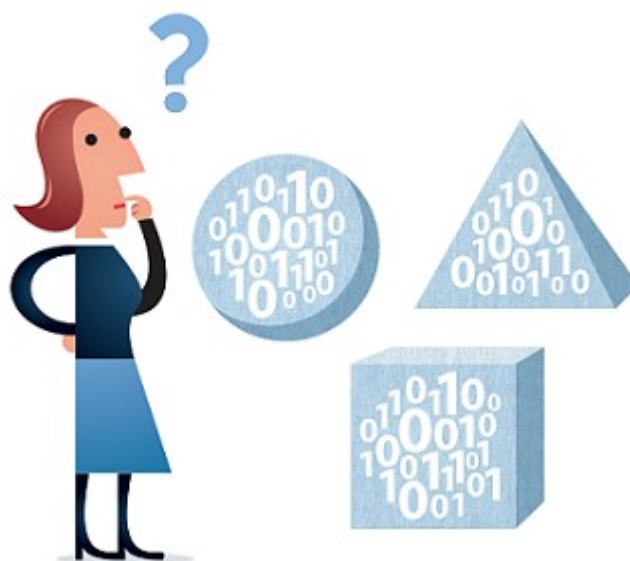


Illustration by Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Introduction

The management of file formats should be considered in the wider strategic context of preservation planning. What can your organisation afford to do? How much developer effort will it require? What do the users require from your collections? Are you committing yourself to a storage problem? At all times, the answer to digital preservation issues is not to try and “do everything”. Your strategy ought to move you towards simple and practical actions, rather than trying to support more file formats than you need.

The purpose of this section is not to provide a detailed or exhaustive list of current formats for different types of content but to draw attention to the broader implications of file formats for their application, and implications for preservation.

A substantial part of this chapter refers to the possible selection of a file format for migration purposes. While migration is a valid preservation strategy, and quite common for many file formats, it is not the only approach or solution. Where appropriate, the chapter will refer to other suitable methods for preservation.

File formats organised by content types

Different content types have, over time, developed their own file formats as they strive to accommodate functionality specific to their needs. The main content types are images, video, audio and text; however, a growing number of formats are being structured to address the demands of new media, including formats for 3D models and archiving the web.

File formats vary enormously in terms of complexity, with some data being encoded in many layers. In some cases the file formats involved are just one part of a larger picture, a picture that includes software, hardware, and even entire information environments.

For further advice on preservation of specific types of digital content and associated file formats see the [Content-specific preservation](#) case studies in the Handbook.

File formats - what should we be worrying about?

Obsolescence

Formats evolve as users and developers identify and incorporate new functionality. New formats, or versions of formats, may introduce file format obsolescence as newer generations of software phase out support for older formats. When software does not provide for backwards compatibility with older file formats, data may become unusable. Both open source and commercial formats are vulnerable to obsolescence: vendors sometimes use planned obsolescence to entice customers to upgrade to new products while open source software communities may withdraw support for older formats if these are no longer generally needed by the community. Obsolescence can also be accidental: both businesses and open source communities can fail.

File format obsolescence is a risk that needs to be understood. That said, the problem may not be as severe as the digital preservation community perceived it to be some 10 years ago. Many established file formats are still with us, still supported, and still usable. It is quite likely that the majority of file formats you deal with will be commonly understood and well supported.

Proliferation

Arguably, in some sectors, proliferation is more of a challenge than obsolescence. If formats aren't normalised then an organisation can end up with a large number of different file formats, and versions of those formats: e.g. lots of different versions of PDF, word, image formats etc. In domains which

develop rapidly evolving bespoke data formats this problem can be exacerbated. Tracking and managing all these formats - which ones are at risk, and which tools can be used for each one - can be a serious challenge.

Your digital preservation strategy should strive to mitigate the effects of obsolescence and proliferation. Strategies as migration, emulation, normalisation and a careful selection of file formats are all valid and worth considering, in the context of your collections and your organisation.

Aspects of file formats for digital preservation

Selecting target formats for preservation

Not all digital formats are suited or indeed designed for archiving or preservation. Any preservation policy should therefore recognise the requirements of the collection content and decide upon a file format which best preserves those qualities. Pairing content with a suitable choice of preservation format or access format; identifying what is important in the content.

Below we suggest some factors to consider in selecting your preferred file formats:

Open source vs proprietary?

Open source formats, such as JPEG2000, are very popular due to their non-proprietary nature and the sense of ownership that stakeholders can attain with their use. However, the choice of open source versus proprietary formats is not that simple and needs to be looked at closely. Proprietary formats, such as TIFF, are seen as being very robust; however, these formats will ultimately be susceptible to upgrade issues and obsolescence if the owner goes out of business or develops a new alternative. Similarly, open source formats can be seen as technologically neutral, being non-reliant on business models for their development however they can also been seen as vulnerable to the susceptibilities of the communities that support them.

Although such non-proprietary formats can be selected for many resource types this is not universally the case. For many new areas and applications, e.g. Geographical Information Systems or Virtual Reality only proprietary formats are available. In such cases a crucial factor will be the export formats supported to allow data to be moved out of (or into) these proprietary environments.

Documentation and standards

The availability of documentation - for example, published specifications - is an important factor in selecting a file format. Documentation may exist in the form of vendor's specifications, an international standard, or may be created and maintained within the context of a user community. Look for a standard which is well-documented and widely implemented. Make sure the standard is listed in the PRONOM file format registry.

Adoption

A file format which is relied upon by a large user group creates many more options for its users. It is worth bearing in mind levels of use and support for formats in the wider world, but also finding out what organisations similar to you are doing and sharing best practice in the selection of formats. Wide adoption of a format can give you more confidence in your preservation strategy.

Lossless vs lossy

Lossy formats are those where data is compressed, or thrown away, as part of the encoding. The MP3 format is widely used for commercial distribution of music files over the web, because the lossy encoding process results in smaller file sizes.

TIFF is one example of an image format that is capable of supporting lossless data. It could hold a high-resolution image. JPEG is an example of a lossy image file format. Its versatility, and small file size, makes it a suitable choice for creating an access copy of an image of smaller size for transmission over a network. It would not be appropriate to store the JPEG image as both the access and archival format because of the irretrievable data loss this would involve.

One rule of thumb could be to choose lossless formats for the creation and storage of "archival masters"; lossy formats should only be used for delivery / access purposes, and not considered to be archival. A rule like this is particularly suitable for a digitisation project, particularly still images.

Support for metadata

Some file formats have support for metadata. This means that some metadata can be inscribed directly into an instance of a file (for example, JPEG2000 supports some rights metadata fields). This can be a consideration, depending on your approach to metadata management.

Significant properties of file formats

This is a complex area. One view regards significant properties as the "essence" of file content; a strategy that gets to the heart of "what to preserve". What does the user community expect from the rendition? What aspects of the original are you trying to preserve? This strategy could mean you don't have to commit to preserving *all* aspects of a file format, only those that have the most meaning and value to the user.

Significant properties may also refer to a very specific range of *technical metadata* that is required to be present in order for a file to be rendered (e.g. image width). Some migration tools may strip out this metadata, or it may become lost through other curation actions in the repository. The preservation strategy needs to prevent this loss happening. It thus becomes important to identify, extract, store and preserve significant properties at early stage of the preservation process.

Things we can do

There are many things you could do to support file formats in your digital archive, and there are many tools available to help you with these tasks. There are now so many that digital preservation tool registries are being developed to help you locate and assess them (see the [Tools](#) and the [Resources](#) sections)

Tools for migration

Broadly, these are tools that transform a file format from an obsolete format into a newer format which can be supported. Many tools exist for doing this migration. They tend to confine themselves to doing one thing (e.g. ImageMagick only works for digital image objects).

A migration tool is just one part of a migration pathway. The pathway must include a destination / target format, which you will have selected in line with guidance as suggested above.

Migration tools may introduce risks. One of these risks is "invisible" changes happening to the content or to the data in the migration. To reduce this risk, one strategy is to devise a set of acceptance criteria for what the transformed object must keep, e.g. in terms of formatting, look and feel, or even functionality, and confirm desired outcomes with a process of quality assurance.

File format migration is not always the solution. Some CAD and CAM file formats cannot easily be migrated, for example. The aerospace industry has found that migration of older CAD files to a newer format requires a lot of validation, mainly because they are required by a regulatory framework to demonstrate that their data is sound and meets very strict standards. In short, the cost of migration and validation is (for them) much higher than an emulation solution, an approach which (in this case) involves keeping the CAD software and maintaining it.

See also the [Tools](#) and [Content-specific preservation](#) sections.

Tools for rendition

Broadly, these are tools that can read and play back a file format, so that the user community can read and interpret the resource; it's most commonly applied to files stored in accessible formats. A basic rendition tool would be PDF Reader. A more sophisticated rendition tool would be the Wellcome Library media player, which supports OCR texts, images, and audio-visual files.

Tools for file format identification

Tools that can identify aspects of file formats which are not immediately obvious from their file extension. They do this by reading the file format header, and thus can identify e.g. mimetype, size, version. Examples of such tools include PRONOM, JHOVE, and the NZ Metadata Extraction Tool (see [Resources](#) below).

These tools are usefully deployed at point of ingest, so that you know from the start what sort of file formats you are taking into the archive.

Some identification tools can also point to likely rendition tools, or even (like PRONOM) suggest a migration path based on file format identification.

Tools for file format validation

JHOVE is one of the few tools that is able to validate a file format. It does this by comparing an instance of a file format with sets of expected behaviours, which it stores in its library. JHOVE can report on certain file formats and tell whether they are valid and well-formed.

Collection surveys

Survey file formats in use / know what you have / characterisation of your collections. This again ties into a planning strategy, letting you know what you need to support, and the likely effort required to do this.

A survey should pay particular attention to *versions* of file formats, and software needed for their reading / rendition. If possible, gather any information about *published specifications* for these formats; some specs are published on the web.

Useful emerging work in this area has taken place at the British Library, with projects on Sustainability Assessments (Maureen Pennock, Paul Wheatley, Peter May) and Collection Profiling (Michael Day, Maureen Pennock, Ann MacDonald). At time of writing there are no active links to these projects, but it is anticipated that the Sustainability Assessment work will be published on the DPC wiki. These are useful approaches and can be regarded as examples of current best practice. Even if you don't assess or profile to the same depth as the BL, the exercise is a practical and applicable one.

Avoid Proliferation of File Types

Where possible, reduce the range of file formats you support, in order to reduce complexity. A sound approach to preservation planning is to normalise, rather than add multiple migration formats to your collection. The smaller the range of formats, the lower the overheads.

Community

Identify a consensus of agreement on target file formats; collaborate with institutions who hold similar collections to yours. What formats do they choose to work with?

Conclusion

For some kinds of content, there is consensus around the choice of preservation format. For example audio archiving where WAV is commonly used. In other areas consensus is much more difficult to achieve. The preservation of digital video is a complex area where progress has been stymied by a lack of agreement, and an uncontrolled proliferation of wrapper formats, delivery methods, and encoding methods. The choice of image file formats is slightly clearer, with a limited choice of formats for archiving and others for delivery. It has been generally agreed that the TIFF format is the correct format for archiving master files (the RAW or DNG format is also considered appropriate for archiving) but this is now being challenged by the JPEG2000 format which provides a far greater level of lossless compression compared to TIFF and is open source.

Resources



Library of Congress recommended format specifications

<http://www.loc.gov/preservation/resources/rfs/index.html>

develop a set of specifications of formats which it recommends, both internally to its own professionals and externally to creators, vendors and archivists, as the preferred ones to use to ensure the preservation and long-term access. It covers both digital and analogue formats and is divided into six broad categories: Textual Works and Musical Compositions; Still Image Works; Audio Works; Moving Image Works; Software and Electronic Gaming and Learning; and Datasets/Databases.

Jisc significant properties reports

Between 2007 and 2008 Jisc funded five studies of significant properties for different types of content and files. Note discussion in the reports is as of 2007- 2008. The reports are as follows:

inSPECT Significant Properties Report 2007 (10 pages)

<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.109.7923&rep=rep1&type=pdf>

Significant Properties of E-learning Objects 2008 (65 pages)

http://www.webarchive.org.uk/wayback/archive/20140616090345/http://www.jisc.ac.uk/media/documents/programmes/preservation/spelos_report.pdf

The Significant Properties of Moving images 2008 (62 pages)

http://www.webarchive.org.uk/wayback/archive/20140616090254/http://www.jisc.ac.uk/media/documents/programmes/preservation/spmovimages_report.pdf

The Significant Properties of Software: A Study 2008 (97 pages)

http://www.webarchive.org.uk/wayback/archive/20100624233431/http://www.jisc.ac.uk/media/documents/programmes/preservation/spsoftware_report_redacted.pdf

The Significant Properties of Vector Images 2007 (61 pages)

http://www.webarchive.org.uk/wayback/archive/20140616090304/http://www.jisc.ac.uk/media/documents/programmes/preservation/vector_images.pdf



British Library File Formats Assessments

http://wiki.dpconline.org/index.php?title=File_Formats_Assessments

The Digital Preservation Team at the British Library has undertaken preservation risk file format assessments to capture knowledge about the gaps in current best practice, understanding and capability in working with specific file formats. The focus of each assessment is on capturing evidence-based preservation risks and the implications of institutional obsolescence which lead to problems maintaining the content over time. The assessments are hosted as a new section on the DPC Wiki. Three assessments covering JP2, TIFF and PDF have commenced the series.

Library of Congress sustainability factors

<http://www.digitalpreservation.gov/formats/index.shtml>

This site is concerned with the formats associated with media-independent digital content, i.e., content that is typically managed as files and which is generally not dependent upon a particular physical medium. It is not concerned with the formats associated with media-dependent digital content, i.e., formats that are dependent upon and inextricably linked to physical media, e.g., DVDs, audio CDs, and videotape formats like DigiBeta. It identifies and describes the formats that are promising for long-term sustainability, and develops strategies for sustaining these formats including recommendations pertaining to the tools and documentation needed for their management.

Jisc digital media infokit: digital file formats

http://www.jiscdigitalmedia.ac.uk/infokit/file_formats/digital-file-formats

This Jisc Digital Media Infokit aims to provide quick and practical answers to 'what file format should I use for...?' It covers still image, audio and video formats and common tasks and applications in education and heritage settings.

Help Solve the File Format Problem

<http://fileformats.archiveteam.org>

A crowd-sourced file format information wiki on the Archive Team site. All content is available under a Creative Commons 0 licence.

Is JPEG 2000 a digital preservation risk?

<http://blogs.loc.gov/digitalpreservation/2013/01/is-jpeg-2000-a-preservation-risk/>

An interesting guest blog and discussion thread on the JPEG 2000 image format.



OPF File Format Risk Registry

<http://wiki.opf-labs.org/display/TR/OPF+File+Format+Risk+Registry>

This focuses specifically on file format issues and risks that have implications for long-term preservation and accessibility and how to deal with these in a practical way. It aims to be complementary to more formal format registries.

PRONOM

<http://apps.nationalarchives.gov.uk/pronom/Default.aspx>

This file format registry is a major resource for anyone requiring impartial and definitive information about the file formats, software products and other technical components required to support long-term access to electronic records and other digital objects of cultural, historical or business value.

DROID (Digital Record Object Identification)

<http://www.nationalarchives.gov.uk/information-management/manage-information/preserving-digital-records/droid/>

This is an automatic file format identification tool providing categories of format identification for unknown files in a digital collection. It uses internal signatures to identify and report the specific file format and version of digital files. These signatures are stored in an XML signature file, generated from information recorded in the PRONOM registry.

Case studies



See the [Detailed content preservation case studies](#) section of the Handbook for relevant case studies.

Information security



Illustration by Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Introduction

This section is intended as guidance for practitioners at a novice or intermediate level on the implications of information security for digital preservation. Information Security issues relate to system security (e.g., protecting digital preservation and networked systems / services from exposure to external / internal threats); collection security (e.g., protecting content from loss or change, the authorisation and audit of repository processes); and the legal and regulatory aspects (e.g. personal or confidential information in the digital material, secure access, redaction). Information security is a complex and important topic for information systems generally. It is important to rely on relevant expertise within your organisation and beyond it through government and other networks for general information security procedures and advice. You may also need appropriate advocacy for specific digital preservation procedures and requirements.

Rigorous security procedures will:

1. Ensure compliance with any legal and regulatory requirements;
2. Protect digital materials from inadvertent or deliberate changes;
3. Provide an audit trail to satisfy accountability requirements;
4. Act as a deterrent to potential internal security breaches;
5. Protect the authenticity of digital materials;
6. Safeguard against theft or loss.

Many types of digital material selected for long-term preservation may contain confidential and sensitive information that must be protected to ensure they are not accessed by non-authorised users. In many cases these may be legal or regulatory obligations on the organisation. These materials must be managed in accordance with the organisation's Information Security Policy to protect against security breaches. ISO 27001 describes the manner in which security procedures can be codified and monitored ([ISO, 2013a](#)). ISO 27002 provides guidelines on the implementation of ISO 27001-compliant security procedures ([ISO, 2013b](#)). Conforming organisations can be externally accredited and validated. In some cases your own organisation's Information Security Policy may also impact on

digital preservation activities and you may need to enlist the support of your Information Governance and ICT teams to facilitate your processes.

Information security methods such as encryption add to the complexity of the preservation process and should be avoided if possible for archival copies. Other security approaches may therefore need to be more rigorously applied for sensitive unencrypted files; these might include restricting access to locked-down terminals in controlled locations (secure rooms), or strong user authentication requirements for remote access. However, these alternative approaches may not always be sufficient or feasible. Encryption may also be present on files that are received on ingest from a depositor, so it is important to be aware of information security options such as encryption, the management of encryption keys, and their implications for digital preservation.

Techniques for protecting information

Several information security techniques may be applied to protect digital material:

Encryption

Encryption is a cryptographic technique which protects digital material by converting it into a scrambled form. Encryption may be applied at many levels, from a single file to an entire disk. Many encryption algorithms exist, each of which scramble information in a different way. These require the use of a key to unscramble the data and convert it back to its original form. The strength of the encryption method is influenced by the key size. For example, 256-bit encryption will be more secure than 128-bit encryption.

It should be noted that encryption is only effective when a third party does not have access to the encryption key in use. A user who has entered the password for an encrypted drive and left their machine powered on and unattended will provide third parties with an opportunity to access data held in the encrypted area, which may result in its release.

Similarly encryption security measures (if used) can lose their effectiveness over time in a repository: there is effectively an arms race between encryption techniques and computational methods to break them. Hence, if used, all encryption by a repository must be actively managed and updated over time to remain secure.

Encrypted digital material can only be accessed over time in a repository if the organisation manages its keys. The loss or destruction of these keys will result in data becoming inaccessible.

Access Control

Access controls allow an administrator to specify who is allowed to access digital material and the type of access that is permitted (for example read only, write). The Handbook follows the National Digital Stewardship Alliance (NDSA) preservation levels in recommending four levels at which digital preservation can be supported through access control. The NDSA levels focus primarily on understanding who has access to content, who can perform what actions on that content and enforcing these access restrictions ([NDSA, 2013](#)) as follows:

NDSA level	Activity
1	<ul style="list-style-type: none"> Identify who has read, write, move and delete authorisation to individual files Restrict who has those authorisations to individual files
2	<ul style="list-style-type: none"> Document access restrictions for content
3	<ul style="list-style-type: none"> Maintain logs of who performed what actions on files, including deletions and preservation actions
4	<ul style="list-style-type: none"> Perform audit of logs

Redaction

Redaction refers to the process of analysing a digital resource, identifying confidential or sensitive information, and removing or replacing it. Common techniques applied include anonymisation and pseudonymisation to remove personally identifiable information, as well as cleaning of authorship information. When related to datasets this is usually carried out by the removal of information while retaining the structure of the record in the version being released. You should always carry out redaction on a copy of the original, never on the original itself.

The majority of digital materials created using office systems, such as Microsoft Office, are stored in proprietary, binary-encoded formats. Binary formats may contain significant information which is not displayed, and its presence may therefore not be apparent. They may incorporate change histories, audit trails, or embedded metadata, by means of which deleted information can be recovered or simple redaction processes otherwise circumvented. Digital materials may be redacted through a combination of information deletion and conversion to a different format. Certain formats, such as plain ASCII text files, contain displayable information only. Conversion to this format will therefore eliminate any information that may be hidden in non-displayable portions of a bit stream.

Resources



ENISA. 2013, Cloud Security Incident Reporting

<https://www.enisa.europa.eu/activities/Resilience-and-CIIP/cloud-computing/incident-reporting-for-cloud-computing/>

The EU's Agency for Network & Information Security offers recommendations on the ways in which cloud providers and their customers should respond to – and report – security breaches. (38 pages).

ISO 27001:2013, Information technology— Security techniques — Information security management systems — Requirements. Geneva: International Organization for Standardization

http://www.iso.org/iso/catalogue_detail?csnumber=54534

ISO 27001 describes the manner in which security procedures can be codified and monitored. Conforming organisations can be externally accredited and validated. A template for a set of policies aligned with the standard is available. Note that these are headings, to assist with policy creation, rather than policy statements. However, similar policy sets are in use in a substantial number of organisations. (23 pages).

ISO 27002:2013, Information technology – Security techniques – Code of practice for information security controls. Geneva: International Organization for Standardization

http://www.iso.org/iso/catalogue_detail?csnumber=54533

ISO 27002 provides guidelines on the implementation of ISO 27001-compliant security procedures. (80 pages)

ISO 27799:2008, Health informatics – Information security management in health using ISO/IEC 27002. Geneva: International Organization for Standardization

http://www.iso.org/iso/catalogue_detail?csnumber=41298

ISO 27799 provides specific advice on implementing ISO 27002 and 27001 in the healthcare sector. (58 pages)



Cabinet Office, 2009, HMG IA Standard No. 1 – Technical Risk Assessment

http://www.cesg.gov.uk/publications/Documents/is1_risk_assessment.pdf

A detailed discussion and standard intended for UK Risk Managers and Information Assurance Practitioners who are responsible for identifying, assessing and treating the technical risks to systems and services that handle, store and process digital government information. (114 pages).

Redaction toolkit (TNA 2011)

http://www.nationalarchives.gov.uk/documents/information-management/redaction_toolkit.pdf

This TNA toolkit was produced in 2011 to provide guidance on editing exempt material from information held by public bodies. It covers generic principles records in any media but has a small section specifically on electronic records and detailed guidance on methods for securely redacting electronic records of all types. (21 pages).

BitCurator

http://wiki.bitcurator.net/index.php?title=Main_Page

BitCurator is a suite of open source digital forensics and data analysis tools to help collecting institutions holding born-digital materials. Parts of the toolset help locate private and sensitive information on digital media and prepare materials for public access.



Information Commissioners Office (ICO): Information security (Principle 7)

<https://ico.org.uk/for-organisations/guide-to-data-protection/principle-7-security/>

The ICO website has guidance on reporting of security breaches and use of IT. For those working in organisations falling under the ICO's jurisdiction an understanding of what this guidance recommends is essential to starting conversations with ICT and Information Governance Colleagues as they will need to be assured that work can be carried out in compliance with ICO recommendations.

Access to the Secure Lab

<http://ukdataservice.ac.uk/get-data/how-to-access/accesssecurelab>

A number of confidential and sensitive microdata sources are becoming available through datalabs across the UK. These data are deemed potentially identifiable, and can only be accessed through a datalab facility (as opposed to download). In addition, researchers are asked to fulfil a number of additional application requirements. Some of these data may be accessed via the Secure Lab of the UK Data Service and this page provides useful overviews and access to relevant user agreements.

Case studies



Opening access to administrative data for evaluating public services: The case of the Justice Data Lab

<http://evi.sagepub.com/content/21/2/232.full.pdf+html>

The Justice Data Lab a unit within a secure setting holding evaluation and statistical expertise has enabled providers of programmes aimed at reducing re-offending to obtain evidence on how the impact of their interventions differs from that of a matched comparison group. This article explores the development of the Justice Data Lab, the methodological and other challenges faced, and the experiences of user organizations. The article draws out implications for future development of Data Labs and the use of administrative data for the evaluation of public services. (16 pages).

UK Data Service: Data Security

<http://ukdataservice.ac.uk/manage-data/store/security.aspx>

This webpage summarises how the UK Data Archive manages data security for its holdings. Data security may be needed to protect intellectual property rights, commercial interests, or to keep sensitive information safe. Arrangements need to be proportionate to the nature of the data and the risks involved. Attention to security is also needed when data are to be destroyed.

References

NDSA, 2013. *The NDSA Levels of Digital Preservation: An Explanation and Uses, version 1* (2013). Available:

http://www.digitalpreservation.gov/ndsa/working_groups/documents/NDSA_Levels_Archiving_2013.pdf

ISO, 2013a. *ISO 27001:2013 - Information technology - Security techniques - Information security management systems - Requirements*. Geneva: International Organization for Standardization. Available: http://www.iso.org/iso/catalogue_detail?csnumber=54534

ISO, 2013b. *ISO 27002:2013 - Information technology – Security techniques – Code of practice for information security controls*. Geneva: International Organization for Standardization. Available: http://www.iso.org/iso/catalogue_detail?csnumber=54533

Cloud services



Illustration by Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

What is cloud computing?

Cloud Computing is a term that encompasses a wide range of use cases and implementation models. In essence, a computing ‘cloud’ is a large shared pool of computing resources including data storage. When someone needs additional computing power, they are simply able to check this out of the pool without much (often any) manual effort on the part of the IT team, which reduces costs and significantly shortens the time needed to start using new computing resources. Most of these ‘clouds’ are run on the public Internet by well-known companies like Amazon and Google. Some larger organisations have also found value in running private clouds inside their own data centres, where similar economies of scale begin to apply.

The generally accepted characteristics of a typical cloud service may be defined as computers and data storage which are:

- Available when required (‘on demand’), without the need for lengthy procurement and configuration processes;

- Available on standard networks such as the Internet, without special requirements for obscure or proprietary networking, protocols, or hardware;
- Able to offer additional capacity as demand increases, and less as demand falls ('elastic');
- Capable of only billing customers for the storage they use.

Cloud computing and digital preservation

Cloud computing can offer several benefits:

- The flexibility of the cloud allows relatively rapid and low-cost testing and piloting of emerging service providers. There are already some pilot activities with these cloud services and opportunities for shared learning across the community;
- There is now much greater flexibility and more options in deployment of cloud storage services and therefore greater relevance to archival repositories compared to earlier years (see Public, Community, Private and Hybrid clouds);
- There are potential cost savings from easier procurement and economies of scale, particularly for smaller repositories. These are important at a time of financial pressures;
- Cloud services can provide easy, automated replication to multiple locations essential for business recovery planning and access to professionally managed digital storage; in addition, the specialists can add access to other dedicated tools, procedures, workflow and service agreements, tailored for digital preservation requirements.

Cloud service models and service providers

There are four different cloud service models:

- Public – Commercial services hosted in large data centres around the world, accessible over public networks to anyone with the means to pay.
- Private - Large organisations create their own cloud by virtualising large sets of physical servers inside their own data centres.
- Hybrid – Combines aspects of combine aspects of public and private cloud , typically to handle large fluctuations in demand, or to satisfy different security requirements.
- Community - Architecturally, it may be effectively the same as a public cloud service, but optimised for a particular group of users to which access is restricted.

There are currently two classes of cloud service provider: generalists offering cloud storage (Amazon, Rackspace, Google, etc), and specialist companies that address additional specific digital preservation requirements and functions (see [Resources and case studies](#) for examples).

Positives

- Cloud services can provide easy, automated replication to multiple locations and access to professionally managed digital storage and integrity checking. As a result bit preservation (durability) of digital information can be at least as good (or better) than can be achieved locally;
- Archives can add access to dedicated tools, procedures, workflow and service agreements, tailored for digital preservation requirements via specialist vendors;

- There are potential cost savings from easier procurement and economies of scale, particularly for smaller archives;
- The flexibility of the cloud allows relatively rapid and low-cost testing and piloting of providers;
- There is much greater flexibility and more options in deployment of cloud services and therefore greater relevance to archives compared to earlier years. In particular private cloud or hybrid cloud implementations can address security concerns over storage of more sensitive material perhaps considered unsuitable for public cloud;
- Exit strategies can be put in place to address archival concerns over provider stability and longevity or other change risks. For example synchronising content across two cloud service providers or an external cloud with local internal storage; or agreeing an escrow copy held independently by a trusted third-party;
- There are already some pilot activities with these cloud services and opportunities for shared learning across the community.

Negatives

- The Cloud is designed for flexibility and rapid change. Archives however are long-term. Cloud storage and service contracts need careful management through time to meet archive needs. Data held in archives must be expected to be both preserved and accessible beyond the commercial lifespan of any current technology or service provider;
- Cloud can be cheaper, but it often requires organisations to think differently about the way their budgets are managed. There are also different skills to IT service vendor and contract management that may involve re-training or recruitment costs;
- Public cloud services tend to bill each month for capacity that has actually been consumed. As a result it can be difficult to budget ahead, or to accurately predict the amount of data likely to be uploaded, stored, or downloaded (however some vendors can invoice you for an annual subscription based on volume);
- As with any form of outsourcing, it is important that archives exercise due diligence in assessing and controlling the risks of cloud storage. Ensure that any legal requirements and obligations relating to third party rights in, or over, the data to be stored will be met. These may relate to management, preservation or access, and may have been placed upon archives and their parent organisations by their donors and funders via contracts and agreements or via legislation by Government;
- Use of cloud services will require archives to consider copyright-related questions including: who currently owns the copyright; whether additional licence permissions may be required; what permissions the cloud provider will need to provide the service; whether the cloud provider is able to use the data for their own purposes; and which party will own the rights in any data or works created from the original data;
- Use of cloud services may raise data security issues, where the relevant data is 'personal data' (e.g. data that permits the identification of a living individual), these include determining responsibility for securing data and audit of providers, as well as about location of processing and the extent to which risks incurred by automation of service provision can be addressed by contract;

- The legal elements of the relationship between an archive and a cloud service provider or providers (e.g. terms of service contracts and service level agreements) must be well defined and meet your requirements. This can be challenging as many cloud providers have standard SLAs and contracts to achieve commodity pricing and have limited flexibility on negotiating terms;
- Explicit provision must be made for pre-defined exit strategies and effective testing, monitoring and audit procedures.

Conclusions

The term "cloud" can encompass a wide range of implementation models for digital preservation services. There is much that can be learnt from organisations who have already piloted or moved to use of the cloud. For example several archives have been able to address the most widely held concerns over cloud services and find ways to successfully integrate cloud storage into their digital preservation activities. Others are using cloud based services for all or part of their other digital preservation functions such as preservation planning. Ultimately, procuring cloud services is similar to procuring any IT. You have to manage and address risks like you would for any other part of your IT infrastructure.

Resources



The National Archives Guidance on Cloud Storage and Digital Preservation (2nd Edition 2015)

http://www.nationalarchives.gov.uk/documents/CloudStorage-Guidance_March-2015.pdf

This guidance explores how cloud storage in digital preservation is developing, emerging options and good practice, together with requirements and standards that archives should consider. Sections focussing on services, legal issues, and five linked case studies, are provided. Sources of further advice and guidance are also included. (39 pages).

Aitken, B, McCann, P, McHugh, A and Miller, K, 2012, Digital Curation and the Cloud, DCC

http://www.jisc.ac.uk/media/7/C/1/%7B7C1A1FD7-44B4-4951-85A8-FC2C4CEB1564%7DCuration-in-the-Cloud_master_final.pdf

This 2012 report focused on the use of cloud services for research data curation. It provides some definitions of Cloud computing and examined a number of cloud approaches open to HE institutions in 2012. (30 pages).

Anderson. S, 2014, Feet On The Ground: A Practical Approach To The Cloud Nine Things To Consider When Assessing Cloud Storage, AV Preserve

<http://www.avpreserve.com/wp-content/uploads/2014/02/AssessingCloudStorage.pdf>

A white paper on cloud services, divided into nine topics and questions to ask. Vendor profiles against these nine topics are available. (7 pages).

A. Brown, C. Fryer, 'Achieving Sustainable Digital Preservation in the Cloud'

<http://www.girona.cat/web/ica2014/ponents/textos/id87.pdf>

This paper describes how Parliament is using the cloud as part of its digital repository infrastructure. 2004 (10 pages).



Digital Preservation Specialist Cloud Service Providers

ArchivesDirect

<http://archivesdirect.org>

ArchivesDirect features a hosted instance of Archivemata with storage via DuraCloud in secure, replicated Amazon S3 and Amazon Glacier storage.

Arkivum

<http://arkivum.com>

Arkivum's Archive as a Service provides a fully-managed and secure service for long-term data retention with online access and a guarantee of data integrity that's part of its Service Level Agreement and backed by worldwide insurance.

DuraCloud

<http://www.duracloud.org>

DuraCloud is a managed service from DuraSpace. It provides support and tools that automatically copies content onto several different cloud storage providers and ensures that all copies of the content remain synchronized. See also ArchivesDirect for its joint service with Archivemata.

Preservica

<http://preservica.com/edition/cloud-edition/>

Preservica Cloud Edition is a fully cloud hosted OAIS compliant digital preservation platform that also includes public access/discovery to allow you to safely share your archive or collection



David Rosenthal's blog

<http://blog.dshr.org/>

Contains a number of posts on the economics of cloud computing

Case studies



The National Archives case study: Archives & Records Council Wales Digital Preservation Working Group

http://www.nationalarchives.gov.uk/documents/Cloud-Storage-casestudy_Wales_2015.pdf

This case study discusses the experience of a cross-sectoral working group of Welsh archives cooperating to test a range of systems and service deployments in a proof of concept for cloud archiving. It explains the organisational context, the varied nature of their digital preservation requirements and approaches, and their experience with selecting, deploying and testing digital preservation in the cloud. The case study examined the open source Archivemata software with Microsoft's Windows Azure; Archivemata with CloudSigma; Preservica Cloud Edition and has begun testing Archivemata with Arkivum 100. January 2015 (10 pages).

The National Archives case study: Tate Gallery

http://www.nationalarchives.gov.uk/documents/Cloud-Storage-casestudy_Tate_Gallery_2015.pdf

This case study discusses the experience of developing a shared digital archive for the Tate's four physical locations powered by a commercial storage system from Arkivum. It explains the organisational context, the nature of their digital preservation requirements and approaches, and their rationale for selecting Arkivum's on-premise solution, "Arkivum/OnSite" in preference to any cloud-based offerings. It concludes with the key lessons learned, and discusses plans for future development. January 2015 (7 pages).

The National Archives case study: Dorset History Centre

http://www.nationalarchives.gov.uk/documents/Cloud-Storage-case-study_Dorset_2015_%281%29.pdf

This case study covers the Dorset History Centre, a local government archive service. It explains the organisational context of the archive, the nature of its digital preservation requirements and approaches, its two year pilot project using Preservica Cloud Edition (a cloud-based digital preservation service), the archive's technical infrastructure, and the business case and funding for the pilot. It concludes with the key lessons they have learnt and future plans. January 2015 (9 pages).

The National Archives case study: Parliamentary Archives

http://www.nationalarchives.gov.uk/documents/Cloud-Storage-casestudy_Parliament_2015.pdf

This case study covers the Parliamentary Archives and their experience of procuring via the G-Cloud framework. For extra resilience/an exit strategy they have selected two cloud service providers with different underlying storage infrastructures. This is an example of an archive using a hybrid set of storage solutions (part-public cloud and part-locally installed) for digital preservation as the archive has a locally installed preservation system (Preservica Enterprise Edition) which is integrated with cloud and local storage and is storing sensitive material locally, not in the cloud. January 2015 (6 pages).

The National Archives case study: Bodleian Library, University of Oxford

http://www.nationalarchives.gov.uk/documents/Cloud-storage-casestudy_Oxford_2015.pdf

This case study covers the Bodleian Library and the University of Oxford, and the provision of a "private cloud" local infrastructure for its digital collections including digitised books, images and multimedia, research data, and catalogues. It explains the organisational context, the nature of its

digital preservation requirements and approaches, its storage services, technical infrastructure, and the business case and funding. It concludes with the key lessons they have learnt and future plans. January 2015 (6 pages).

King's College London Kindura Project

<http://link.springer.com/article/10.1186%2F2192-113X-2-13>

The Kindura project led by King's College London and funded by Jisc, sought to pilot the use of a hybrid cloud for research data management. It used DuraCloud to broker between storage or compute resources supplied by external cloud services, shared services, or in-house services. There is an earlier Jisc prepared case study but a more recent open-access article on the project is linked.

University of Illinois Archives 2011 evaluation of Archivemata

<http://e-records.chrisprom.com/evaluating-open-source-digital-preservation-systems-a-case-study-2/>

Angela Jordan describes a 2011 evaluation by the University of Illinois Archives of Archivemata—an open-source, OAIS Reference Model-compliant digital preservation system. Because Archivemata was then in its alpha stages, working with this system was a way to explore what the system offered in relation to the needs of the University Archives, as well as provide input to the developers as they continued to refine the software for production release.

Digital forensics



Illustration by Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Introduction

Digital forensics is associated in many people's minds primarily with the investigation of wrongdoing. However, it has also emerged in recent years as a promising source of tools and approaches for facilitating digital preservation and curation, specifically for protecting and investigating evidence from the past.

Institutional repositories and professionals with responsibilities for personal archives and other digital collections can benefit from forensics in addressing digital authenticity, accountability and accessibility. Digital personal information must be handled with due sensitivity and security while demonstrably protecting its evidential value.

Forensic technology makes it possible to: identify privacy issues; establish a chain of custody for provenance; employ write protection for capture and transfer; and detect forgery or manipulation. It can extract and mine relevant metadata and content; enable efficient indexing and searching by curators; and facilitate audit control and granular access privileges. Advancing capabilities promise increasingly effective automation in the handling of ever higher volumes of personal digital information. With the right policies in place, the judicious use of forensic technologies will continue to offer theoretical models, practical solutions and analytical insights.

Forensics in practice

There are three basic and essential principles in digital forensics: that the evidence is acquired without altering it; that this is demonstrably so; and that analysis is conducted in an accountable and repeatable way. Digital forensic processes, hardware and software have been designed to ensure compliance with these requirements.

Information assurance is critical. Writeblockers ensure that information is captured without altering it, while chains of custody in terms of evidence handling, process control, information audit, digital signatures and watermarking protect the historical evidence from future alteration and uncertain provenance.

Selective redaction, anonymization and encryption, malware sandbox containment and other mechanisms for security and fine-tuned control are required to assure that privacy is fully protected and inadvertent information leakage is prevented. Family computers, portable devices and shareable cloud services all harbour considerable personal information and consequently raise issues of privacy. Digital archivists and forensic practitioners share the need to handle the ensuing personal information responsibly.

The current emphasis on automation in digital forensic research is of particular significance to the curation of cultural heritage, where this capability is increasingly essential in a digital universe that continues to expand exponentially. Current research is directed at handling large volumes efficiently and effectively using a variety of analytical techniques. Parallel processing, for example, through purpose-designed Graphics Processing Units (GPUs), and high performance computing can assist processor-intensive activities such as full search and indexing, filtering and hashing, secure deletion, mining, fusion and visualization.

Especially noteworthy for digital preservation and curation is the way that digital forensics directs attention towards the digital media item as a whole – typically the forensic disk image, the file that represents everything on the original disk.

Forensic technologies

Forensic technologies vary greatly in their capability, cost and complexity. Some equipment is expensive, but some is free. Some techniques are very straightforward to use, others have to be applied with great care and sophistication. The BitCurator Consortium has been an important development bringing together a community of archival users of open source digital forensic tools ([Lee et al, 2014](#)). There is an increasingly rich set of open source forensic tools that are free to obtain and use – most significantly for archivists, BitCurator. These are a wonderful introduction to the ins-and-outs of digital forensics, and can be used to compare and cross-check the outputs of commercial or other open source tools.

Digital archivists and forensic specialists share a common need to monitor and understand how technology is used to create, store, and manage digital information. Additionally, there is a mutual need to manage that information responsibly in conformance with relevant standards and best

practice. New forensic techniques are furthering the handling of digital information from mobile devices, networks, live data on remote computers, flash media, virtual machines, cloud services, and encrypted sources. The use of encryption is beginning to present significant challenges for digital preservation. It is not only a matter of decryption but of identifying encryption in the first place. Digital forensics offers some solutions.

Forensic and archival methodology must retain the ability both to retrospectively interpret events represented on digital devices, and to react quickly to the changing digital landscape by the rapid institution of certifiable and responsible policies, procedures and facilities. The pace of change also has implications for ongoing training of curators and archivists, and there are digital forensics courses endorsed by archival, scholarly and preservation institutions.

Conclusion

In conclusion, there are some deep challenges ahead for cultural heritage and archives, but the forensic perspective is undoubtedly among the most promising sources of insights and solutions. Equally, digital forensics can benefit from the advances being made in the curation and preservation of digital information.

This brief overview has been based on short excerpts from The Digital Preservation Technology Watch Report on Digital Forensics and Preservation ([John, 2012](#)) with additional material kindly provided by Jeremy Leighton John, the author of the report. See [Resources and case studies](#) for further detailed guidance and exemplars.

Resources



Digital forensics and preservation DPC technology watch report

<http://dx.doi.org/10.7207/twr12-03>

This 2012 DPC report provides a broad overview of digital forensics, with some pointers to resources and tools that may benefit cultural heritage and specifically the curation of personal digital archives (60 pages).

Digital forensics and born-digital content in cultural heritage collections

<http://www.clir.org/pubs/reports/pub149/pub149.pdf/view>

This CLIR report introduces the field of digital forensics in the cultural heritage sector and explores some points of convergence between the interests of those charged with collecting and maintaining born-digital cultural heritage materials and those charged with collecting and maintaining legal evidence (93 pages).



Archivematica

https://www.archivemata.org/wiki/Main_Page

Archivemata is an open source digital preservation system and has addressed the ingest of forensic disk images as part of its workflows and toolset.



BitCurator

<http://www.bitcurator.net>

The website provides access to information on the BitCurator Consortium (BCC), projects, and tools. BitCurator has developed, packaged and documented open-source digital forensics tools to allow libraries, archives and museums to extract digital materials from removable media in ways that reflect the metadata and ensure the integrity of the materials, allowing users to make sense of materials and understand their context, and preventing inadvertent disclosure of sensitive data. The consortium is an independent, community-led membership association that serves as the host and center of administrative, user and community support for the BitCurator environment.

Forensics wiki

http://forensicswiki.org/wiki/Main_Page

The Forensics Wiki is a Creative Commons-licensed wiki devoted to information about digital forensics. It lists over 700 pages focused on the tools and techniques used by investigators, important papers and reports, people, and organizations involved.



The Invisible Photograph Part 2: Trapped: Andy Warhol's Amiga Experiments

<http://www.nowseethis.org/invisiblephoto/posts/108>

A team of computer scientists, archivists, artists, and curators teamed up to unearth Andy Warhol's lost digital works on a 30 year old Amiga Commodore computer (18 mins 52 secs)

The Invisible Photograph Part 3: Extraterrestrial: The Lunar Orbiter Image Recovery Project

<http://www.nowseethis.org/invisiblephoto/posts/384>

How the "techno archaeologists" of the Lunar Orbiter Image Recovery Project digitally recovered the first photographs of the moon taken by a set of unmanned space probes in the 1960s. (22 mins 07 secs)

Case studies



Carcenet email project

<http://www.library.manchester.ac.uk/aboutus/projects/carcenet/>

A Jisc-funded project that tackled the challenge of capturing and preserving the email archive of Carcanet Press, one of the UK's premier poetry publishing houses. It was winner of the 2014 DPC Preservation Award for Safeguarding the Digital Legacy. The project explored and adopted several ediscovery and forensic tools, specifically AccessData's Forensic Toolkit (FTK), Paraben's Email Examiner and Fookes Software's Aid4Mail eDiscovery. A project final report summarizes the work ([Baker, 2014](#)).

References

John, J. L., 2012. Digital Forensics and Preservation. *DPC Technology Watch Report 12-03* November 2012. Available: <http://dx.doi.org/10.7207/twr12-03>

Lee, C. A., Olsen, P., Chassanoff, A., Woods, K., Kirschenbaum, M. & Misra, S., 2014. *From Code to Community: Building and Sustaining BitCurator through Community Engagement*. BitCurator White Paper 30 September 2014. Available: <http://www.bitcurator.net/wp-content/uploads/2014/11/code-to-community.pdf>

Persistent identifiers

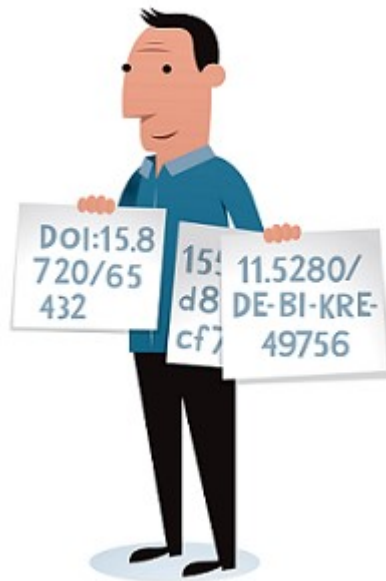


Illustration by Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Introduction

This section provides guidance on the use of persistent identifiers for digital objects and digital preservation. Other types of persistent identifier schemes exist e.g. for individuals or institutions.

A persistent identifier is a long-lasting reference to a digital resource. Typically it has two components: a unique identifier; and a service that locates the resource over time even when it's location changes. The first helps to ensure the provenance of a digital resource (that it is what it purports to be), whilst the second will ensure that the identifier resolves to the correct current location.

Persistent identifiers thus aim to solve the problem of the persistence of accessing cited resource, particularly in the academic literature. All too often, web addresses (links) fail to take you to the referenced resource you expect. This can be for technological reasons like server failure but human-created failures are more common. Organisations transfer journals to new publishers, reorganise their websites, or lose interest in older content, leading to broken links when you try to access a resource. This is frustrating for users, but the consequences can be serious if the linked resource is essential for legal, medical or scientific reasons.

Persistent identifiers can also be used 'behind-the-scenes' within a repository to manage some of the challenges in cataloguing and describing, or providing intellectual control and access to born-digital materials.

Schemes

Since the problem of persistence of an identifier is created by humans, the solution of persistent identifiers also has to involve people and services not just technologies. There are several persistent identifier schemes and all require a human service element to maintain their resolution systems. The main persistent identifier schemes currently in use are detailed below.

Digital Object Identifier (DOI)

[DOIs](#) are digital identifiers for objects (whether digital, physical or abstract) which can be assigned by organisations in membership of one of the DOI Registration Agencies; the two best known ones are CrossRef, for journal articles and some other scholarly publications, and DataCite for a wide range of data objects. As well as the object identifier, DOI has a system infrastructure to ensure a URL resolves to the correct location for that object.

Handle

[Handles](#) are unique and persistent identifiers for Internet resources, with a central registry to resolve URLs to the current location. Each Handle identifies a single resource, and the organisation which created or now maintains the resource. The Handle system also underpins the technical infrastructure of DOIs, which are a special type of Handles.

Archival Resource Key (ARK)

[ARK](#) is an identifier scheme conceived by the California Digital Library (CDL), aiming to identify objects in a persistent way. The scheme was designed on the basis that persistence "is purely a matter of service and is neither inherent in an object nor conferred on it by a particular naming syntax".

Persistent Uniform Resource Locator (PURL)

[PURLs](#) are URLs which redirect to the location of the requested web resource using standard HTTP status codes. A PURL is thus a permanent web address which contains the command to redirect to another page, one which can change over time.

Universal Resource Name (URN)

[URNs](#) are persistent, location-independent identifiers, allowing the simple mapping of namespaces into a single URN namespace. The existence of such a Uniform Resource Identifier does not imply availability of the identified resource, but such URIs are required to remain globally unique and persistent, even when the resource ceases to exist or becomes unavailable. The URN term is now deprecated except in the very narrow sense of a formal namespace for expressing a Uniform Resource Identifier.

Choosing a Persistent Identifier Scheme

There needs to be a social contract to maintain the persistence of the resolution service - either by the organisation hosting the digital resource, a trusted third party or a combination of the two. Each scheme has its own advantages and constraints but it is worth considering the following when deciding on a persistent identifier strategy or approach:

Advantages

- Critically important in helping to establish the authenticity of a resource.
- Provides access to a resource even if its location changes.
- Overcomes the problems caused by the impermanent nature of URLs.
- Allows interoperability between collections.

Disadvantages

- There is no single system accepted by all, though DOIs are very well established and widely deployed.
- There may be costs to establishing or using a resolver service.
- Dependence on ongoing maintenance of the permanent identifier system.

Conclusions

Persistent identifiers need to be supported by enduring services and are not just unique strings of alpha-numeric characters that are assigned to a digital resource. They have become particularly important for research data and e-journal articles (see content specific preservation section on [e-Journals](#)) and are a significant part of the long-term infrastructure for digital preservation of research. For the issue of link-rot for more general web pages, and solutions harnessing web-archives to resolve this see the content specific preservation section on [Web-archiving](#).

Resources



Persistent identifiers - an overview. TWR Technology Watch Review

<http://www.metadaten-twr.org/2010/10/13/persistent-identifiers-an-overview/>

This article by Juha Hakala (2010) describes five persistent identifier systems (ARK, DOI, PURL, URN and XRI) and compares their functionality against the cool URIs. The aim is to provide an overview, not to give any kind of ranking of these systems.

Preservation, trust and continuing access for e-Journals DPC technology watch report

<http://dx.doi.org/10.7207/twr13-04>

This 2013 report by Neil Beagrie discusses current developments and issues which libraries, publishers, intermediaries and service providers are facing in the area of digital preservation, trust and continuing access for e-journals. It includes generic lessons and recommendations on outsourcing and trust of

interest to the wider digital preservation community and covers relevant legal, economic and service issues as well as technology. (49 pages).

Persistent Identifiers in the Publication and Citation of Scientific Data

http://www.iza.org/en/papers/5090_28102009.pdf

Presentation by Jens Klump, German Research Centre for Geosciences (GFZ) on the DFG STD-DOI project, which details the background and reasoning behind the foundation of DataCite. 2009. (47 pages).

DCC Briefing Paper: Persistent Identifiers

<http://www.dcc.ac.uk/resources/briefing-papers/introduction-curation/persistent-identifiers>

This 2006 paper by Joy Davidson discusses how progress in defining the nature and functional requirements for identifier systems is hindered by a lack of shared agreement on what identifiers should actually do; simply provide a globally or locally unique name for a digital or analogue resource, or incorporate associated services such as resolution and metadata binding. The application and maintenance of identifiers forms just one part of an overall digital preservation strategy; in order to offer any guarantees of persistence in the long or short-term they need institutional commitment and clearly defined roles and responsibilities. (2 pages)



ARK

<http://www.cdlib.org/services/uc3/arkspec.pdf>

CrossRef

<http://www.crossref.org>

DataCite

<http://www.datacite.org>

DOI

<http://www.doi.org/>

Handle

<http://www.handle.net/>

Perma.CC

<https://perma.cc/about>

PURL

<https://purl.org/docs/index.html>

URN

<http://tools.ietf.org/html/rfc3986>

Case studies



DCC case study: Assigning digital object identifiers to research data at the University of Bristol

<http://www.dcc.ac.uk/resources/persistent-identifiers>

The University of Bristol runs a dedicated research data repository as part of their Research Data Service. They are using the DataCite service at the British Library to assign digital object identifiers (DOIs) to research datasets in order to provide unique and perpetual identifiers for data, to allow easy citation and discoverability. The Bristol Research Data Service provides guidance on how to use the identifiers to cite data and is developing appropriate policies to monitor usage. 2004. (4 pages).

Links that Last

<http://www.dpconline.org/events/previous-events/925-links-that-last>

This DPC briefing day in July 2012 introduced the topics of persistent identifiers and linked data, and discussed the practical implications of both approaches to digital preservation. It considered the viability of services that offer persistent identifiers and what these offer in the context of preservation; reviewed recent developments in linked data, considering how such data sets might be preserved; and by introducing these two parallel topics it went on to consider whether both approaches can feasibly be linked to create a new class of robust linked data. A series of presentations including case studies are linked from the provisional programme.