



Préservation du contenu de l'Audio visuel numérique

Les enregistrements de l'audio visuel (AV) du vingtième siècle sont en danger, et la numérisation a été une solution qui a créé un nouveau problème : la préservation des contenus des AV numériques. Ces fichiers ont des caractéristiques (taille, formats spécifiques) qui ne sont pas bien pris en compte par la technologie actuelle. Les « pratiques les plus modernes » peuvent être recommandées, mais trois changements majeurs sont nécessaires : 1) Les technologies existantes de la préservation numérique et des silos de stockage devraient être utilisées pour les collections d'av ; 2) La technologie devrait évoluer afin de prendre en compte les média dépendant du temps ; 3) Les technologies de l'information en général et des supports de masse devraient aussi évoluer afin de rendre en compte les besoins spécifiques des fichiers audio visuels.

Les problèmes de la préservation de l'audio visuel numérique

L'existant européen (dans les archives et autres collections officielles) est estimé à 50 millions d'heures de vidéo, de films, d'audio, la plupart en formats analogiques. Environ 70 pour cent de ces objets sont en danger maintenant, et l'ensemble le sera dans moins de trente ans – à cause des dégradations diverses, des détériorations et de l'obsolescence des formats.

D'importants programmes de numérisation ont démarré : dans la dernière décade, environ 10 millions d'heures ont été numérisées. Pendant que l'on s'occupait de changer en fichiers les contenus des cassettes, bandes magnétiques et disques – en tant que solution de préservation – le reste du monde avait pris conscience que les fichiers numériques posaient leurs propres problèmes de préservation.

Les grosses collections de fichiers numériques posent un problème de management technique ; la solution réside dans la technologie des silos de stockage. Les fichiers ont besoins d'être maintenus : ils doivent être nommés, déplacés sur de nouveaux supports (fréquemment), copiés pour des demandes d'accès, encodés pour des besoins d'accès divers, être soumis à des contrôles de validité. Ils nécessitent des interventions sur leurs métadonnées, allant du catalogage au robot moissonneur (pour les accès globaux et normalisés). Une maintenance manuelle est simplement impossible – et comporte trop de risques d'erreurs – dès que les collections atteignent une certaine taille. La technologie des silos de stockage fournit des outils automatisés pour la création, la maintenance et les besoins de manipulation et d'accès pour les grands ensembles de fichiers. Il existe de nombreuses publications sur cette technologie.

Deux mondes : La technologie des silos de stockage (bibliothèque numérique) émane du monde des bibliothèques académiques. Les collections d'AV sont majoritairement en dehors de ce monde. Les plus grands détenteurs de contenus sont les diffuseurs ; d'autres grands détenteurs sont des musées du film et autres institutions culturelles et patrimoniales (une des plus grande collection de films en Grande Bretagne se trouve au musée impérial de la guerre).

Les diffuseurs peuvent être différents, mais il n'est pas rare que les équipes informatiques d'un diffuseur, ainsi que l'encadrement qui décide et finance les solutions techniques, ne connaissent absolument rien du monde des bibliothèques académiques et des technologies des silos de stockage.

Le premier obstacle que l'on rencontre en voulant préserver les fichiers AV est de connaître, de comprendre, de financer et d'utiliser les outils existants de cette technologie des silos de stockage, qui peuvent transformer une montagne de fichiers en une collection gérable.

Le second obstacle est de reconnaître que les outils des silos de stockage permettent la gestion (ainsi les fichiers peuvent être accédés et ne se perdent pas) mais ne couvrent pas la préservation. Les fichiers sont soumis à différents problèmes d'obsolescence qui sont pris en compte par la technologie de la préservation numérique - méthodes pour s'assurer que les fichiers en voie d'obsolescence peuvent être migrés vers de nouvelles normes et de nouveaux formats, méthodes pour émuler des environnements de vieux système d'exploitation afin d'étendre la durée d'utilisation de formats devenant obsolètes, des critères pour évaluer la fiabilité d'un dépôt numérique, et enfin une méthodologie globale : OAIS.

Il est difficile de trouver quelqu'un, dans les équipes informatiques s'occupant des collections d'AV, qui ait entendu parler d'OAIS, ce qui limite singulièrement l'aide au financement et à l'implémentation.

Heureusement, le projet de la Communauté européenne Memories développe OAIS et les processus qui y sont reliés, spécialement pour les collections audio et vidéo.

Le troisième obstacle est que les besoins particuliers des fichiers AV ne sont pas complètement supportés ni par les technologies des silos de stockage ni par celles de la préservation des objets numériques, comme discuté plus loin.

Références et informations complémentaires

[1] Etudes sur la situation de l'audiovisuel:

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

PrestoSPACE: <http://www.prestospace.eu>

TAPE: <http://www.tapeonline.net/survey.html>

http://www.tapeonline.net/docs/audiovisual_research_collections.pdf

[2] Guides généraux de la préservation:

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

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

[3] Une liste des outils de la conservation numérique

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

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

DROID:

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

PRONOM:

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

National Library of New Zealand Library metadata

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

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

OAIS:

http://www.dpconline.org/docs/lavoie_OAIS.pdf

Migration:

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

Emulation:

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

10granger.html

Repository evaluation criteria reference:

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

[4] Projets et initiatives

MEMORIES:

<http://www.memories-project.eu/>

European

Digital Library:

<http://www.europeana.eu/>

Les problèmes particuliers aux données et aux fichiers numériques de l'audiovisuel

A cause du problème « des deux mondes », les formats de diffusion professionnels (MXF en particulier) ne sont pas supportés par la plupart des outils de préservation et des silos de stockage. D'autres formats « standards » sont mieux supportés mais beaucoup (ex AVI, WMV) sont propriétaires, ce qui en soit même est un problème pour la préservation.

Les problèmes restants sont liés aux contenus actuels des fichiers.

- La plupart des fichiers AV sont compressés. Quelle que soit la 'qualité originelle' qui a été perdue lors de la compression, elle l'est pour toujours. La préservation devrait privilégier la rétention de la qualité ; une possibilité qui a besoin d'être définie puis incorporée à la technologie actuelle.

- Le contenu lié au temporel nécessite des outils dotés d'une dimension temporelle (catalogage, navigation, modification ...).

- Les fichiers sont complexes. En effet, le concept d'enveloppe a été défini pour prendre en compte cette complexité d'un fichier AV typique : signaux multiples, divers types de métadonnées, y compris les domaines du temporel (sous titrage) et des valeurs chiffrées (valeurs de durées).

- La préservation audiovisuelle doit prendre en compte beaucoup de fichiers interdépendants : encodages avec et sans perte, multiples logiciels d'interface (qui supportent les accès en différents formats, par exemple Real, Windows Média, MPEG, AVI, Quicktime, Flash), différents paliers de modification et de recomposition, et une série de données sur les droits : les différentes parties prenantes, les différents organismes de perception, la non uniformité de pays à pays. Un grand ensemble d'informations, décrivant les signaux, les métadonnées et les droits, doit être préservé.

L'accès

Les bibliothèques ont une longue tradition de l'accès uniformisé : des catalogues communs reposant sur des métadonnées normalisées, afin de fournir le service 'n'importe quel livre n'importe où'. Beaucoup de collections audiovisuelles ont pour tradition d'être fermées, ou ouvertes uniquement pour des accès professionnels ou commerciaux.

Les bibliothèques numériques perpétuent cette tradition de l'accès uniformisé et étendu, souvent à une échelle nationale ou internationale comme la Bibliothèque Numérique Européenne. Les collections audio visuelles ont besoin de la technologie des silos de stockage pour être accessibles à travers les principaux projets, tel celui de la BNE. En retour, l'industrie des silos de stockage doit faire plus d'efforts pour prendre en compte les problèmes des données et des fichiers numériques de l'audio visuel, tels qu'exposés plus haut.

Plus particulièrement, les silos de stockage ont besoin d'outils pour les accès liés au temporel pour à la fois le signal audio visuel et les métadonnées (les droits, par exemple, peuvent être différents, d'un moment à l'autre, dans un même fichier audio visuel).

La plupart des contenus audio visuels sont détenus par des organismes qui n'ont pas pour coutume de travailler avec des bibliothèques et qui préfèrent limiter l'accès à 'leurs' contenus.

Les problèmes de marchés, de promotion et de droits gênent un portail de l'audiovisuel européen. La BNE n'inclura peut être jamais les contenus de la BBC.

Que faire

En dépit des problèmes, des conclusions claires peuvent être tirées au sujet de la préservation audio visuelle :

- Préserver l'objet : Garder 'l'original', même s'il est compressé. 'Préserver la chaîne de bits' quoi que vous fassiez d'autre. Le contenu audio visuel a un avantage : il est très volumineux, mais en un relativement petit nombre de formats. Des méthodes pour 'traiter/ jouer les bits' existent peut être.

- Décoder sans compression puis sauvegarder toujours sans compresser (en plus de conserver l'original). Ceci est une obligation contraignante pour la vidéo (100 GB /heure pour la télévision 625 lignes), mais le stockage est maintenant peu onéreux.

- Renforcer les métadonnées : l'extension de fichiers (ex .wav, .avi) n'est pas suffisante. Il existe plus de 50 variantes d'encodage pour la définition de .wav ; MPEG-1 et MPEG-2 sont sous l'extension .mpg. Au mieux, il y aura un outil d'extraction des métadonnées ; sinon, on aura recours à la documentation manuelle et aux tests manuels.

- Vous n'êtes pas seul : utilisez les centres d'enregistrement des types de fichiers, les entrepôts de progiciels, les plates formes d'émulation, et les publications citées en références sur la préservation.

AUTEUR: Richard Wright, *Research and Innovation, BBC*

Richard.wright@bbc.co.uk

TRADUCTEUR: Jean-Pierre Hazel Teil