



Paris, le 21 mai 2010

Le directeur, chargé des archives de France

à

Mesdames les directrices des services à compétence nationale des Archives nationales

Mesdames et Messieurs les directeurs des Archives départementales

sous couvert de Mesdames et Messieurs les préfets

Archives de France
Sous-direction de la politique interministérielle et territoriale
des archives traditionnelles et numériques

Michel Jacobson
01 40 27 62 64

Note d'information DGP/SIAF/2010/010 relative à la diffusion d'un guide méthodologique pour le choix de formats numériques pérennes dans un contexte adaptés aux données orales et audiovisuelles.

J'ai le plaisir de vous annoncer la parution d'un guide méthodologique pour le choix de formats numériques pérennes favorisant la conservation des données orales et audiovisuelles.

Ce guide est le produit d'un travail qui a été confié par les Archives de France au département « archivage diffusion » du Centre informatique national de l'enseignement supérieur (CINES), dans le cadre d'une recherche financée par le CNRS en lien avec le Très Grand Equipement (TGE) Adonis.

Le choix des formats et des encodages est une question cruciale dans la mesure où c'est la capacité à maîtriser parfaitement le format des fichiers pris en charge pour leur préservation dans le temps, qui permet ou non d'assurer la lisibilité et l'intelligibilité de l'information.

De même, le contrôle des formats au moment de l'introduction dans les systèmes d'archivage est une étape essentielle car c'est de la connaissance et de la maîtrise de la structure des données que dépend le succès des opérations d'interprétation et de transformation. Comme nombre des formats récents sont de simples conteneurs, ce contrôle doit porter, non pas uniquement sur l'enveloppe, mais aussi sur les codages des contenus.

Les migrations de formats, rendues indispensables en cas d'obsolescence, sont des étapes critiques dans le processus de pérennisation car la capacité à conserver le contenu des informations ainsi que les fonctionnalités permises par le format d'origine dépend des choix effectués. Des compromis doivent souvent être trouvés entre le poids du fichier et les pertes admises à la suite de compressions.

Comme il existe des centaines de formats et familles de formats qui évoluent très rapidement, des études devaient être menées afin de déterminer leur capacité à représenter l'information d'une manière pérenne, et afin d'identifier ou de développer les outils nécessaires à leur identification, à leur contrôle, voire à leurs conversions. L'ampleur de ce travail et la somme des compétences qu'il réclame, rendaient nécessaire de mutualiser ces études entre partenaires scientifiques institutionnels, chacun suivant son domaine de spécialisation.

Le CINES a ainsi mis à disposition sur les réseaux l'outil Facile¹ qu'il a développé et qui permet aux internautes d'analyser le degré de conformité à un format d'un fichier dont on souhaite connaître les caractéristiques. Cet outil fournit la liste des formats acceptés pour archivage sur la plateforme d'archivage électronique du CINES, ainsi que les logiciels d'identification et de contrôle utilisés. La vérification permet de déterminer si le fichier est valide et bien formé par rapport au format déclaré².

Les premiers formats audiovisuels qui ont été étudiés par le CINES s'inscrivent dans le cadre d'un projet-pilote sur l'archivage pérenne de corpus de données orales (enregistrements de chercheurs étudiant les langues et la parole). Le guide présente le résultat synthétique de cette étude.

Dans la première partie sont indiquées les caractéristiques des formats numériques puis celles propres aux formats audiovisuels. Sont ensuite explicités les critères retenus pour la sélection des formats audiovisuels garantissant un archivage pérenne.

La seconde partie du guide fournit une liste (non exhaustive) de formats audiovisuels disponibles actuellement sur le marché, avant de détailler ceux qui ont été retenus dans le cadre du projet d'archivage pérenne des données orales du TGE-Adonis. Enfin une annexe analyse sous forme de fiche structurée (20 paramètres), tous les formats étudiés.

Le directeur

Hervé LEMOINE

1 FACILE validation du Format d'Archivage du Cines par anaLyse et Expertise, à l'adresse suivante :<http://www.cines.fr/FACILE-validation-du-Format-d.html>

2 Un fichier est considéré comme "bien formé" s'il répond aux exigences purement **syntaxiques** de la spécification de son format. Un fichier est considéré comme "valide" s'il répond aux exigences **sémantiques** de la spécification de son format.

