


Boîte à outils d'archivage pour une bonne gouvernance

DIRECTIVE 18 :
Préservation numérique





La version originale de cette directive a été préparée par la Branche Régionale pour le Pacifique du Conseil International d'Archives (PARBICA), pour être utilisée par les pays du Pacifique. Cela signifie que la directive peut faire référence à des éléments que vous ne maîtrisez pas ou qui ne sont pas utilisés dans votre pays. Vous devrez probablement modifier certains des conseils de cette directive pour les adapter aux dispositions de votre propre gouvernement. Pour obtenir une copie modifiable de cette directive, merci de contacter les archives nationales, le Bureau des archives publiques ou toute autre autorité en charge des archives dans votre pays, ou contactez PARBICA à l'adresse suivante : <http://www.parbica.org>.

Boîte à outils d'archivage pour une bonne gouvernance

Directive 18 : Préservation numérique

TABLE DES MATIÈRES

À qui est destinée cette directive ?	4
Les cinq règles d'or de la préservation numérique.....	5
Préservation de l'accès numérique.....	6
Préservation de l'authenticité numérique.....	8
Stratégies de préservation numérique.....	9
Migration des documents d'activité.....	11
Approche de la préservation numérique en 8 étapes.....	13
Glossaire.....	15
Autres documents à consulter	17
Annexe a - formats de fichier - ouverts et propriétaires	18
Annexe b - outils de préservation numérique.....	21
Annexe c : dix étapes pour conserver des documents numériques exploitables.....	24
Annexe d : reformatage numérique d'enregistrements audiovisuels analogiques....	27

À QUI EST DESTINÉE CETTE DIRECTIVE ?

Cette directive est essentiellement destinée au personnel des archives et des bibliothèques des îles du Pacifique, dont l'activité de base consiste à préserver l'accès à l'information sur le long terme. Cette directive vise à fournir à ce personnel des conseils pratiques sur les questions associées à la préservation numérique et sur la façon d'aborder ces questions. Un public secondaire est le personnel des ministères, en particulier le personnel en charge de la gestion des documents d'activité et de l'information et le personnel en charge des technologies de l'information. Les membres du public secondaire doivent veiller à lire l'annexe C de cette directive, intitulée « Comment éviter la perte d'informations à l'ère du numérique », qui est conçue comme un document séparé destiné au personnel des ministères.

Cette directive est structurée pour passer d'une discussion sur les motifs expliquant l'importance de la préservation numérique à un contenu plus technique sur les formats de fichiers et les outils de préservation numérique, en passant par une présentation des approches et des stratégies de préservation numérique. La directive se termine par un glossaire, une liste de lectures complémentaires, deux annexes techniques et une annexe qui résume les messages clés destinés aux responsables et au personnel des départements gouvernementaux et portant sur la façon d'éviter la perte d'informations à l'ère du numérique. La dernière annexe, l'annexe D, bien qu'elle ne concerne pas la préservation des documents d'activité numériques à la création, traite des questions et des approches du reformatage numérique des matériaux audiovisuels analogiques à des fins de préservation.

LES CINQ RÈGLES D'OR DE LA PRÉSERVATION NUMÉRIQUE

1. Gérez vos documents d'activité numériques comme vos documents papier : utilisez des systèmes de classement structurés, un bon titrage des dossiers et de bonnes règles de nommage.
2. Ne conservez que les documents d'activité numériques que vous devez conserver (en d'autres termes, sachez ce que vous devez conserver et supprimez tout le reste dès qu'il n'est plus nécessaire).
3. Lorsque vous entreprenez des actions de préservation, ne travaillez que sur une seule copie.
4. Sauvegardez toujours les choses.
5. Faites-vous des amis parmi vos collègues des services informatiques.

PRÉSERVATION DE L'ACCÈS NUMÉRIQUE

Comme tous les documents d'activité papier, les documents d'activité numériques doivent être préservés de façon à pouvoir être retrouvés, consultés et utilisés aussi longtemps qu'ils sont nécessaires.

La préservation numérique comprend l'ensemble des activités menées pour préserver l'information numérique. Les informations numériques peuvent inclure une série de documents d'activité numériques à la création (créés à l'origine dans un format numérique), tels qu'e-mails, pages web, documents et feuilles de calcul, ainsi que des documents numériques repris qui ont été produits à partir de matériel analogique dans le cadre d'un projet de numérisation ou de conversion/reformatage. La préservation numérique ne fait pas référence au processus de copie de matériel analogique vers une forme numérique (connu sous le nom de numérisation), mais plutôt aux actions entreprises sur les fichiers numériques créés du fait de la numérisation.

La nature complexe et fragile des documents d'activité numériques signifie qu'il existe un risque important que les données qu'ils contiennent soient inaccessibles, même après seulement quelques années. Les données numériques ont en moyenne un cycle de vie de seulement sept à dix ans.

Pour consulter un document d'activité papier, il suffit d'avoir le document (ou d'en avoir une copie) et de comprendre la langue dans laquelle il est écrit. Les documents d'activité numériques sont rédigés en langage informatique, qui est incompréhensible pour les personnes. En d'autres termes, les documents d'activité numériques dépendent de la technologie. Pour compliquer encore les choses, les données du fichier sont écrites de différentes manières (appelées formats de fichier), en fonction du programme qui les crée. Les détails exacts du format de fichier ne sont souvent connus que du fabricant/vendeur du programme qui le crée. Pour accéder aux données stockées dans un document d'activité numérique, vous avez besoin d'un logiciel capable d'interpréter le format du fichier et de le représenter sur un écran d'ordinateur de façon compréhensible.

La préservation des objets numériques doit être plus proactive que celle des documents papier. Si vous pouvez ranger un livre sur une étagère et y revenir 100 ans plus tard en vous contenant de l'ouvrir pour le lire, utiliser la même approche pour un objet numérique garantit presque que vous ne pourrez pas y accéder dans le futur.

Obsolésence technologique

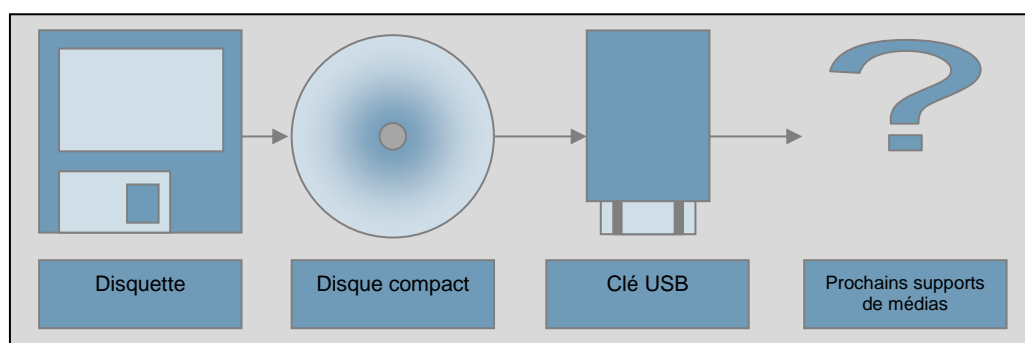
L'un des principaux défis associés à la préservation des documents d'activité numériques consiste à tenir compte du fait que la technologie informatique est rapidement obsolète. C'est ce qu'on appelle l'obsolésence technologique. Lorsqu'un logiciel particulier est remplacé ou qu'une nouvelle version de celui-ci apparaît, ou lorsqu'un dispositif matériel n'est plus produit par l'industrie, les documents créés avec ces technologies peuvent ne plus être accessibles. Des stratégies de préservation sont alors nécessaires pour garantir que les documents d'activité numériques vivent au-delà de la durée de vie du système avec lequel ils ont été créés. Pour garantir la continuité de l'information numérique, il est essentiel que cette information soit parfaitement gérée dès sa création.

Exemple d'obsolésence technologique

Vous avez créé un document sur un vieil ordinateur de travail il y a des années, et vous l'avez enregistré sur un lecteur de disquettes de 5 ¼ pouces. Vous souhaitez maintenant consulter ce document mais votre lieu de travail ne dispose plus d'ordinateurs équipés de lecteurs de disquettes.

Vous contactez un ami qui a le bon lecteur sur son ordinateur. Vous insérez une disquette et pouvez voir votre document, mais vous ne pouvez pas ouvrir le fichier, car l'application logicielle que vous avez utilisée pour créer le fichier n'est plus utilisée, et il n'existe aucun autre programme qui comprenne ce format de fichier.

Ce qui aurait pu être fait : Une solution au problème de l'obsolésence matérielle aurait consisté à migrer continuellement le fichier vers des supports actuels (un CD par exemple) avant que votre lieu de travail se débarrasse des ordinateurs équipés de lecteurs de disquettes.



Pour éviter le problème de l'obsolésence des logiciels, vous auriez pu faire migrer le contenu d'un format de fichier numérique vers un autre à mesure des mises à jour.

PRÉSERVATION DE L'AUTHENTICITÉ NUMÉRIQUE

L'un des principaux défis de la préservation numérique consiste à garantir l'authenticité des documents d'activité au fil du temps. Comme les documents d'activité sont une source de preuve des décisions et des activités, il est essentiel d'être assuré que les documents d'activité sont effectivement ce qu'ils prétendent être, et que les utilisateurs puissent être sûrs que les documents n'ont pas été altérés, corrompus ou modifiés d'une quelconque façon. Lorsque les documents d'activité sont dignes de confiance et fiables en tant que preuves, qu'ils peuvent être examinés par un tribunal, ils sont authentiques. La facilité avec laquelle les documents d'activité numériques peuvent être modifiés, que ce soit délibérément ou accidentellement, fait de l'authenticité un sujet essentiel. Alors qu'un document d'activité papier contient certaines caractéristiques physiques qui sont des indicateurs évidents d'authenticité (comme un en-tête ou une signature), ces caractéristiques ne sont pas toujours aussi évidentes ou fixes sur un document d'activité numérique.

Lorsqu'un document d'activité est déplacé d'un système à un autre (comme lors d'une mise à niveau du logiciel), ou d'un support à un autre (s'il est copié d'un serveur à un CD, par exemple), il risque de perdre son intégrité et son authenticité par corruption des fichiers. L'actualisation des données, la mise en œuvre de bonnes pratiques informatiques pour la sécurité des données et les sauvegardes, et la conservation de copies multiples des flux binaires originaux sont autant de pratiques permettant de garantir l'authenticité des objets numériques.

STRATÉGIES DE PRÉSERVATION NUMÉRIQUE

Comme pour les stratégies de préservation du papier, il n'existe pas de stratégie unique pour la préservation numérique. Il peut être nécessaire d'utiliser une série de stratégies adaptées aux différentes catégories de matériel numérique.

Les approches courantes de la préservation numérique sont les suivantes :

1. La préservation de la technologie - cette approche implique la préservation du logiciel original qui a été utilisé pour créer et accéder à l'information, et la préservation du système d'exploitation original et du matériel sur lequel il fonctionne. Cette stratégie n'est pas considérée comme réalisable sur le long terme, bien qu'il puisse être utile de conserver une ou deux générations de logiciels et/ou de matériels antérieurs.
2. Émulation - cette approche consiste à émuler d'anciens logiciels et matériel en utilisant des ordinateurs modernes. L'émulation nécessite des ressources et une expertise technique importantes, et n'est généralement pas une option viable pour les organisations de petite taille ou de taille moyenne.
3. Migration - cette approche consiste à transformer l'information numérique en nouveaux formats avant que l'ancien format ne devienne obsolète, tout en préservant le contenu intellectuel de l'information et en conservant la capacité des utilisateurs à récupérer, afficher et utiliser l'information.

Parmi ces trois approches, la migration est la stratégie de préservation numérique la plus fiable, la plus rentable et la plus couramment utilisée. Elle est expliquée dans le détail aux pages 9 et 10 qui suivent.

Copies de sauvegarde des informations numériques

Il est très important de réaliser des copies de sauvegarde des informations numériques de votre organisation afin d'éviter la perte de ces informations en cas de catastrophe. La sauvegarde consiste à copier des informations sur différents supports de stockage, tels qu'un disque dur séparé ou des bandes. *Conservez toujours au moins deux copies des informations, une copie de travail et une copie de sauvegarde, de préférence hors site.*

Les copies de sauvegarde sont très importantes pour garantir la poursuite des activités en cas d'urgence, mais on ne peut pas s'y fier en tant que système de tenue d'archives ou stratégie d'archivage sur le long terme. Par exemple, les copies de sauvegarde ne comportent généralement pas de métadonnées, et les données peuvent être sauvegardées en une masse unique, ce qui rend la localisation et la récupération de fichiers spécifiques très difficiles. Il existe également le risque que les données sauvegardées ne soient pas accessibles au fil du temps sans le logiciel adéquat. Dans la mesure du possible, exportez les métadonnées des référentiels numériques et des systèmes de SGEDD et incluez-les dans la sauvegarde. Les disques durs externes sont très bon marché de nos jours, et la mise en œuvre d'un système de copies de sauvegarde hors site simple est relativement facile.

Le service informatique d'une organisation est généralement responsable de la réalisation des copies de sauvegarde des informations numériques. Les gestionnaires des documents d'activité doivent s'entretenir avec leur service informatique pour garantir que les documents d'activité et documents sur support de sauvegarde remplissent leur fonction et sont conservés aussi longtemps que nécessaire.

Gestion des supports de stockage

Les supports physiques utilisés pour stocker les données (bandes magnétiques, disques durs, etc.) sont beaucoup plus vulnérables aux dommages physiques que les livres et autres objets en papier. Si le papier est susceptible de se détériorer ou d'être attaqué par des moisissures ou des insectes, la détérioration est lente et peut ne pas se manifester avant plusieurs décennies. Pour les objets numériques en revanche, les dommages peuvent être rapides et instantanés. Une petite rayure sur un CD peut rendre immédiatement inaccessibles tous les fichiers qu'il contient ; un disque dur externe ou un ordinateur portable peut vous échapper des mains et être endommagé irrémédiablement en une seconde.

À mesure que la quantité de documents numériques créés et diffusés augmente, la gamme et la variété des supports de stockage mobile utilisés pour stocker et partager des informations numériques augmentent également. Les clés USB, les CD et les DVD ne sont que quelques exemples de supports qui peuvent être utilisés pour stocker des informations à des fins de sauvegarde ou de mobilité. Les supports de stockage mobile ne doivent pas être utilisés pour conserver les copies originales des documents, et doivent être conservés dans un endroit sûr lorsqu'ils ne sont pas utilisés.

MIGRATION DES DOCUMENTS D'ACTIVITÉ

La migration doit être :

1. planifiée, en évaluant et en gérant les risques.
2. réalisée aussi vite que possible, car attendre peut être plus coûteux, voire rendre la migration impossible.
3. réalisée avec soin.
4. surveillée et documentée.
5. vérifiée, pour garantir que les informations ont conservé leur sens, leur utilité et leur intégrité.

La migration est le transfert de données ou d'une ressource numérique vers un environnement d'exploitation plus récent ou vers un environnement moins susceptible d'être affecté par l'obsolescence technologique. Il peut s'agir de la conversion d'un format de fichier à un autre (par exemple la conversion de Microsoft Word au format OpenDocument), d'un système d'exploitation à un autre (par exemple de Windows à Linux) ou d'un langage de programmation à un autre (par exemple de C à Java) afin que les données restent pleinement accessibles et fonctionnelles.

La migration vers de nouveaux environnements d'exploitation signifie souvent que la copie n'est pas exactement la même que l'élément d'information original. Certaines fonctionnalités, telles que les macros dans une feuille de calcul Excel, peuvent être perdues lors de la migration entre différents formats de fichiers (voir l'annexe A pour plus d'informations sur les formats de fichiers). Lorsque la migration entraîne une perte d'informations ou de fonctionnalités, on parle de « migration avec perte ». La migration avec perte est souvent plus problématique lorsque le fichier d'origine est dans un format propriétaire, car les outils peuvent ne pas être en mesure d'interpréter correctement la structure des données. Les risques liés à la lecture de données dans des formats de fichiers propriétaires ne font qu'empirer avec le temps. Il est donc généralement préférable de migrer le plus tôt possible pour contribuer à atténuer ce risque.

Avant d'entreprendre une migration, il est important d'évaluer la quantité d'informations perdues au cours du processus et de procéder à une analyse des risques pour déterminer si les conséquences de cette perte d'informations sont insignifiantes ou significatives. Si la perte d'informations est identifiée comme significative, le choix d'une autre stratégie peut être justifié. Quoi qu'il en soit, ces délibérations doivent être documentées à des fins de référence future et à des fins d'audit, notamment pour permettre à l'organisation de justifier de l'authenticité des informations par rapport aux traitements de préservation qui ont été utilisés.

La migration peut se faire de différentes façons. Dans sa forme la plus simple, la migration peut impliquer la copie des informations numériques sur un support non numérique plus stable, par exemple sur du papier ou un microfilm. Si le papier et les microfilms peuvent être plus fiables sur le long terme, il arrive souvent qu'ils ne ressemblent pas complètement à l'objet numérique original ou qu'ils n'aient pas les mêmes fonctionnalités que celui-ci. Le transfert des fichiers vers une version plus récente du support de stockage, ou vers un support de stockage différent (par exemple, d'une disquette à un CD ou à un disque dur externe, par fois appelé « rafraîchissement du support ») offre une stratégie à court ou moyen terme pour préserver l'accès, mais nécessite toujours la migration des fichiers numériques lorsque la technologie change.

Une autre approche pour préserver l'accès aux informations numériques consiste à les faire migrer dès le début vers des formats standard, qui peuvent être plus stables que les formats originaux. Le choix d'un format de préservation des informations numériques dépendra de l'aspect de la ressource qui sera nécessaire ultérieurement. Par exemple, la nécessité de traiter ou d'éditer une ressource numérique à l'avenir ou de préserver une présentation visuelle aura un impact sur le format utilisé pour préserver l'information numérique.

La complexité du processus de migration variera en fonction de la ressource numérique que l'on fait migrer. La migration d'un objet multimédia interactif sera beaucoup plus complexe que la migration d'un simple texte. Dans sa forme la plus complexe, la migration peut prendre du temps et être coûteuse, même si elle sera toujours beaucoup moins chère que de tenter dans le futur de lire des données obsolètes. Dans tous les cas, la migration doit être planifiée et budgétisée comme une partie intégrante de la maintenance continue des systèmes et des données.

Il convient de ne pas oublier que, comme la gestion des documents d'activité numériques en général, la préservation numérique entraîne des coûts permanents. Les coûts comprennent la main-d'œuvre, la maintenance du matériel et des logiciels, le remplacement des supports et des biens d'équipement, les licences logicielles, les coûts de stockage et les coûts d'électricité.

APPROCHE DE LA PRÉSERVATION NUMÉRIQUE EN 8 ÉTAPES

Lorsqu'une organisation s'engage à garantir la préservation sur le long terme de ses informations numériques, l'approche suivante en 8 étapes vous aidera à mettre en œuvre une stratégie de préservation numérique.

Étape	Activité	Exemples
1	Planifier	<ul style="list-style-type: none">• Lier les activités de préservation à l'évaluation des documents d'activité pour garantir que vous ne gaspillez pas vos efforts en préservant des documents d'activité ayant une valeur sur le court terme.• Intégrer les activités de préservation dans la planification normale des programmes de gestion des documents d'activité et des archives.• Effectuer un contrôle régulier des documents d'activité numériques.• Identifier et tester les outils de préservation numérique que vous souhaitez utiliser.
2	Établir des partenariats et communiquer avec vos partenaires	Avec : <ul style="list-style-type: none">• Personnel des TIC, pour leur savoir-faire technique.• Direction, pour l'affectation des ressources.• Experts externes, autorités archivistiques et groupes du secteur de la préservation numérique.
3	Conserver les documents d'activité dans des systèmes capables de garantir l'intégrité, l'authenticité et l'accessibilité des documents d'activité	<ul style="list-style-type: none">• Les documents d'activité doivent être contrôlés et identifiés pour permettre un suivi de l'accessibilité.• Les documents d'activité doivent être fixés de façon à ce qu'ils ne puissent pas être modifiés ou altérés.
4	Utiliser de bonnes métadonnées de tenue d'archives	Tenir à jour et associer aux documents d'activité des informations sur : <ul style="list-style-type: none">• le contexte opérationnel dans lequel les documents d'activité ont été créés.• leurs dépendances techniques, toute activité de migration ou de conversion.• les caractéristiques essentielles qui préservent et garantissent l'authenticité, par exemple une carte peut avoir besoin d'être conservée en couleur.

Étape	Activité	Exemples
5	Utiliser des formats ouverts	<ul style="list-style-type: none"> • Dans la mesure du possible, utiliser des formats de données ouverts pour faciliter la migration. • Consulter l'Annexe A pour plus d'informations.
6	Gérer les supports de stockage	<ul style="list-style-type: none"> • Effectuer des contrôles de routine des supports de stockage. • Remplacer régulièrement les supports afin de limiter le risque d'endommagement ou de perte des documents d'activité. • Stocker les supports dans un environnement stable et contrôlé.
7	Migrer - déplacer les documents d'activité vers de nouveaux formats, supports et systèmes	<ul style="list-style-type: none"> • Conserver des documents d'activité sur des supports de stockage récents, par exemple sur un disque dur externe plutôt que sur une disquette de 5¼ pouces. • Faire évoluer les documents d'activité à mesure que les versions du logiciel changent. • Planifier, gérer et documenter chaque migration.
8	Protéger les actifs numériques	<ul style="list-style-type: none"> • Effectuer des sauvegardes régulières à des fins de reprise après sinistre. • Veiller à ce que les documents d'activité soient protégés contre les virus et les pirates informatiques. • Utiliser une protection par mot de passe et d'autres contrôles d'accès.

GLOSSAIRE

Authenticité – un document d'activité est considéré comme authentique s'il peut être prouvé

- qu'il est ce qu'il dit être ;
- qu'il a été créé ou envoyé par la personne supposée l'avoir créé ou envoyé ;
et
- qu'il a été créé ou envoyé à la date indiquée.

Somme de contrôle - nombre obtenu via l'exécution d'un algorithme à partir du contenu d'un document numérique et destiné à vérifier l'exactitude de la transmission, de la copie ou de l'enregistrement du contenu informatif d'un document numérique.

Fichier numérique - assemblage logique de données binaires lisibles par machine, stocké dans un système informatique.

Signature numérique – informations qui, à l'aide de techniques cryptographiques, fournissent des garanties quant à l'authenticité et/ou la fiabilité et/ou la paternité d'un fichier numérique.

Émulation - utilisation de programmes qui imitent le matériel et le logiciel originaux (obsolètes ou non disponibles) utilisés pour créer un fichier numérique.

Format de fichier - façon particulière dont les informations sont codées pour être stockées dans un fichier informatique et pour être utilisées par une application logicielle.


Migration avec perte - se produit lorsque la migration de fichiers informatiques entraîne la perte d'informations ou de fonctionnalités dans ces fichiers.

Métadonnées – données décrivant le contexte, le contenu et la structure des documents d'activité et qui permettent leur localisation, leur utilisation, leur gestion et leur conservation dans le temps.

Migration – modification du format des informations numériques pour que celles-ci puissent être visualisées et utilisées avec des matériels et des logiciels différents, généralement plus récents ou à plus long terme, tout en maintenant l'authenticité, l'intégrité, la fiabilité et la facilité d'utilisation de ces informations.

Logiciel libre – logiciel informatique distribué gratuitement dans le cadre d'un accord de licence et dont le code informatique peut être partagé, visualisé et modifié par d'autres utilisateurs et organisations. Le développement de logiciels à code source ouvert est souvent réalisé par une communauté distribuée de développeurs de logiciels via l'Internet.

Format propriétaire – applications logicielles et/ou formats de fichiers développés, détenus et contrôlés par une entité commerciale privée, dont le code logiciel ou les spécifications ne sont pas facilement accessibles et ne peuvent généralement pas être utilisés sans payer de droits de licence.



Rafraîchir - copier des informations numériques d'un support de stockage à un autre pour protéger les informations contre toute perte causée par la détérioration du support de stockage d'origine.

Support de stockage - support physique des informations numériques, dont des exemples sont les disques durs, les bandes numériques, les disques durs externes, les lecteurs flash, les disques compacts, les disquettes.

Virus - programme ou code informatique qui est transféré dans un système informatique ou un fichier numérique à l'insu de l'utilisateur et dans l'intention de corrompre ou de supprimer des informations dans l'ordinateur destinataire.

AUTRES DOCUMENTS À CONSULTER

Brown, Adrian, *Archiving Websites : A practical guide for information management professionals* (Sites Web d'archivage : guide pratique à destination des professionnels de la gestion des informations), Facet Publishing, Londres, 2006.

Digital Preservation Coalition, *Preservation Management of Digital Materials : The Handbook* (Gestion de la préservation des matériels numériques : le manuel) York, 2008, <http://www.dpconline.org/advice/digital-preservation-handbook.html>

Fleischhauer, Carl, « Format considerations in audiovisual preservation reformatting: Snapshots from the Federal Agencies Digitization Guidelines Initiative » (Considérations sur le format lors du reformatage pour préservation audiovisuelle : instantanés de l'initiative Directives de numérisation pour les agences fédérales) *Information Standards Quarterly*, 22 (2), printemps 2010, p. 34 - 40, http://www.digitizationguidelines.gov/audio-visual/documents/IP_Fleischhauer_AudioVisual_Reformatting_isqv22no2.pdf

Harvey, Ross, *Digital Curation : A how-to-do-it manual* (Organisation de contenu : manuel pratique) Facet Publishing, Londres, 2010.

Harvey, Ross, *Preserving Digital Materials* (Préserver les matériaux numériques), K.G. Saur, Munich, 2005.

Association internationale des archives sonores et audiovisuelles, *Guidelines on the Production and Preservation of Digital Audio Objects* (Directives sur la production et la préservation des objets audio numériques), 2^e édition, 2009, <http://www.iasa-web.org/tc04/audio-preservation>

Jasper, Mike, « How to convert reel to reel audio to digital » (Comment convertir de l'audio de bobine à bobine en numérique), http://www.ehow.com/how_5796206_convert-reel-reel-audio-digital.html

Lawrence, Alexis, « How to convert old VHS tapes to DVDs » (Comment convertir de vieilles cassettes VHS en DVD), http://www.ehow.com/how_6743273_convert-old-vhs-tapes-dvds.html

Müller, Heiko, *Database Archiving* (Archivage des bases de données), Digital Curation Centre, Edinburgh, 2009, http://www.era.lib.ed.ac.uk/bitstream/1842/3346/3/Mueller%20Database%20archiving%20_%20Briefing%20Paper.doc

RLG-NARA Task Force on Digital Repository Certification, *Trustworthy Repositories Audit and Certification : Criteria and Checklist* (Audit et certification de référentiels sécurisés : critères et liste de contrôle), Center for Research Libraries, Chicago, 2007, http://www.crl.edu/sites/default/files/attachments/pages/trac_0.pdf

ANNEXE A - FORMATS DE FICHER - OUVERTS ET PROPRIÉTAIRES

Un format de fichier correspond à la façon particulière dont les informations sont codées pour être stockées dans un fichier informatique et utilisées par une application logicielle. Les formats de fichier sont indiqués par une extension qui apparaît à la fin du titre du document ou de l'objet numérique, par exemple « .html », « .pdf » ou « .doc ». Le choix des formats de fichiers à utiliser pour créer des documents d'activité numériques à la création doit être fait en tenant compte de la durabilité à long terme, ainsi que des exigences opérationnelles immédiates. Il est utile d'identifier un ensemble minimum de formats qui répondent à la fois aux besoins opérationnels actifs et aux critères de durabilité, et de ne créer les données que dans ces formats.

Il existe deux grands types de formats de fichiers, les *formats ouverts* et les *formats propriétaires*. Pour utiliser un format propriétaire, il est nécessaire d'accepter (et souvent de payer) les conditions de licence imposées par la société privée qui possède le code source de ce format. Les formats ouverts ne sont pas détenus par des propriétaires privés, mais sont détenus et développés par une communauté, bien qu'il vous faille toujours accepter de respecter des conditions de licence pour utiliser des formats ouverts. Les formats ouverts sont également appelés formats de fichiers libres s'ils ne sont pas liés par des restrictions telles que des droits d'auteur, des brevets et des marques commerciales, si bien que chacun peut les utiliser gratuitement à toutes fins utiles, sous réserve de certaines conditions d'utilisation. Si les formats ouverts sont généralement préférables à des fins de préservation numérique, un format ouvert n'est pas nécessairement un bon format. Un bon indicateur de la qualité d'un format ouvert est le degré d'adoption de ce format par les organisations et les membres de la communauté.

Les formats ouverts et les formats propriétaires sont expliqués plus avant par la suite.

	Format ouvert	Format propriétaire ouvert	Format propriétaire fermé
Exemples	Portable Network Graphics (PNG), OpenDocument Format (ODF), JPEG, Extensible Hypertext Markup Language (XHTML), Tagged Image File Format (TIFF), Free Lossless Audio Codec (FLAC), Portable Document Format (PDF) - ISO 32000.	MP3 audio, Microsoft Office Open XML – par exemple Docx (OoXML, également publié comme norme ECMA 376 et ISO 29500 transitoire).	Microsoft Word (DOC), Microsoft Outlook, Excel (XLS), PowerPoint (PPT), Photoshop (PSD), Microsoft Access, formats d'image RAW, format audio WAV.
Qu'est-ce que c'est ?	Propriété intellectuelle partagée publiquement, généralement entretenue par un organisme de normalisation.	Propriété intellectuelle privée.	Propriété intellectuelle privée.
Disponibilité de la spécification de format	Spécification publiée sans aucune restriction.	La spécification peut être mise à disposition avec restriction.	Aucune spécification accessible au public.
Comment est-il développé ?	Le format est élaboré dans le cadre d'un processus visible par le public et mené par la communauté.	Le format est développé et commercialisé par des entreprises qui contrôlent la façon dont la technologie est utilisée, afin d'améliorer leur position sur le marché.	Le format est développé et commercialisé par des entreprises qui contrôlent la façon dont la technologie est utilisée, afin d'améliorer leur position sur le marché.
Comment l'utiliser ?	Peut être utilisé et modifié par quiconque sans restriction, à l'exception des conditions de licence qui peuvent limiter le développement de versions commercialisées des logiciels.	Le titulaire de la licence a le contrôle exclusif de la technologie à l'exclusion des autres (actuels ou futurs).	Le titulaire de la licence a le contrôle exclusif de la technologie à l'exclusion des autres (actuels ou futurs).

	Format ouvert	Format propriétaire ouvert	Format propriétaire fermé
Quel logiciel est nécessaire pour l'utiliser ?	Les formats ouverts peuvent être mis en œuvre par toute personne, y compris des logiciels propriétaires et des logiciels libres et gratuits, en utilisant les licences typiques utilisées par chacun.	En général, seules des applications sous licence peuvent utiliser librement ces formats.	Les formats de fichiers propriétaires ne sont accessibles qu'à l'aide du logiciel qui a produit ce fichier, ou des applications sous licence.
Conclusion	Il est recommandé d'utiliser si possible les formats ouverts, qui sont pris en charge par un large éventail de logiciels ou qui sont indépendants d'une plate-forme. Les formats ouverts favorisent la viabilité sur le long terme des données en permettant la migration d'un environnement technique à un autre, sans dépendre d'un fournisseur spécifique.	Les formats propriétaires ouverts présentent un risque plus élevé que les formats ouverts, car ils sont contrôlés par une société dans le cadre d'accords de licence qui peuvent changer.	Les formats propriétaires présentent un plus grand risque concernant l'accessibilité des données qu'ils contiennent sur le long terme. Le manque de documentation sur la spécification et les exigences en matière de licence des logiciels signifient que le format est moins à l'épreuve du temps.

ANNEXE B - OUTILS DE PRÉSERVATION NUMÉRIQUE

De nombreuses solutions permettent de collecter, d'identifier et de gérer les données, dont un certain nombre d'outils gratuits et à code source ouvert. Toutes les solutions nécessitent - au minimum - un poste de travail informatique, un membre du personnel ayant une expérience en informatique pour mettre en œuvre le logiciel et importer et gérer les données, et un espace de stockage des données (deux disques durs au minimum).

Nom	Que fait-il ?	Qui pourrait l'utiliser ?
Outils de préservation numérique à code source ouvert		
<p>HOPPLA (Home and Office Painless Persistent Long-Term Archiving)</p> <p>HOPPLA est gratuit et peut être téléchargé gratuitement à l'adresse suivante : http://www.ifs.tuwien.ac.at/dp/hoppla/release/index.html#anchor3.</p> <p>Publié en septembre 2010</p>	<p>HOPPLA propose un ensemble de stratégies permettant la préservation des données numériques, y compris la réalisation de copies de sauvegarde et la migration. HOPPLA copie des fichiers à différents endroits et convertit certaines de ces copies dans d'autres formats s'il existe un risque que le fichier source original devienne obsolète. L'outil peut traiter des documents, des photos et/ou vidéos personnelles et des collections de musique numérique.</p>	<p>Utilisateurs privés et petites organisations qui savent qu'ils ont besoin d'un système de préservation numérique mais qui n'ont pas les compétences ou les connaissances nécessaires pour en développer un.</p>
<p>Plate-forme logicielle de préservation numérique des Archives nationales d'Australie (DPSP)</p> <p>Le code source et la documentation de soutien au système DPSP sont gratuits et disponibles à l'adresse suivante : http://dpsp.sourceforge.net/.</p>	<p>La DPSP permet de faire migrer les données de leur format d'origine vers des formats ouverts et entièrement documentés à des fins de conservation des archives. Les outils DPSP comprennent Manifest Maker (qui crée une liste de fichiers numériques et la dirige vers un emplacement spécifique), XENA (qui convertit les fichiers dans des normes de format ouvertes), Digital Preservation Recorder (l'outil de flux de travail qui ingère et stocke les fichiers) et Checksum Checker (qui surveille que le contenu des archives numériques ne soit pas corrompu). Le DPSP est un système de référentiel qui s'interface avec la base de données de gestion des collections des Archives nationales, et utilise le code-barres unique appliqué au dossier par l'organisation.</p>	<p>Organisations de grande taille, disposant de ressources importantes et ayant accès à des compétences professionnelles et à une bonne infrastructure en matière de TIC.</p>

Nom	Que fait-il ?	Qui pourrait l'utiliser ?
Archivemata http://archivemata.org/wiki/index.php?title=Software	Archivemata est un système différent de la DPSP, qui repose sur des référentiels. Au lieu d'utiliser une base de données complexe ou un système de contrôle externe qui suit les liens entre l'objet et ses métadonnées, il utilise un système dans lequel l'objet et les métadonnées restent dans le même système de fichiers.	La simplicité d'Archivemata le rend idéal pour les organisations dont les capacités financières et techniques sont limitées.
Outils de plate-forme ou de cadre (applications logicielles qui fonctionnent comme des sous-ensembles d'une archive numérique)		
Fedora Commons (Flexible Extensible Digital Object and Repository Architecture)	Fedora fournit une architecture de gestion des actifs numériques (DAM), sur laquelle de nombreux types de référentiels institutionnels, d'archives numériques et de systèmes de bibliothèques numériques peuvent être construits. Il s'agit d'un logiciel libre largement utilisé, qui peut être exécuté sur différents systèmes d'exploitation, mais dont la mise en œuvre adéquate exige une grande expertise technique.	Organisations de grande taille, disposant de ressources importantes et ayant accès à des compétences professionnelles et à une bonne infrastructure en matière de TIC.
DSpace	DSpace fournit les outils nécessaires à la gestion des ressources numériques et une plateforme pour les activités de préservation. Il s'agit également d'un logiciel libre, qui offre un système de référentiel pouvant gérer un large éventail d'objets numériques, qui peut être exécuté sur différents systèmes d'exploitation et dispose d'une large base d'utilisateurs, en particulier dans les établissements d'enseignement, mais aussi moins fréquemment dans des agences gouvernementales.	Établissements d'enseignement qui doivent gérer de grandes quantités de ressources numériques.
Outils d'archivage web - la technique d'archivage web la plus courante utilise des « robots d'indexation » - un programme informatique qui navigue sur le web pour automatiser le processus de collecte des pages web.		
Robot d'indexation Heritrix (développé par l'Internet Archive, une importante bibliothèque numérique)	Prend en charge les processus clés tels que les autorisations, la planification des tâches, la récolte, l'examen de la qualité et la collecte de métadonnées descriptives.	

Nom	Que fait-il ?	Qui pourrait l'utiliser ?
Web Curator Tool (WCT)	Une application open-source de gestion de flux de travail pour l'archivage sélectif du web. Il est conçu pour être utilisé dans les bibliothèques et d'autres organisations de collecte, et prend en charge la collecte par des utilisateurs non techniques tout en permettant un contrôle complet du processus de moissonnage du web. Il utilise le robot d'indexation Heritrix.	
HT Track	Permet de télécharger un site Web depuis Internet vers un répertoire local, en organisant la structure des liens relatifs du site d'origine. L'utilisateur peut ouvrir une page du site web « en miroir » dans son navigateur et parcourir le site de lien en lien, comme s'il le visualisait en ligne.	
Outils de migration libres		
Image Magick www.imagemagick.org	Outil pour la migration des images.	
SoX www.sox.sourceforge.net	Outil pour la migration du son.	

TRAC

Les organisations qui créent de nouveaux référentiels numériques et qui souhaitent évaluer la solidité et la durabilité de leurs référentiels peuvent utiliser TRAC (Trusted Repositories Audit and Certification: Criteria and Checklist) (audits et certification de référentiels de confiance : critères et liste de contrôle).¹ En général, TRAC :

- Fournit des outils permettant l'audit, l'évaluation et la certification éventuelle de référentiels numériques
- Établit les exigences de documentation requises pour l'audit
- Définit un processus de certification.

¹ Online Computer Library Centre (OCLC) et Center for Research Libraries (CRL), *Trusted Repositories Audit and Certification: Criteria and Checklist*, 2007, http://www.crl.edu/sites/default/files/attachments/pages/trac_0.pdf

ANNEXE C : DIX ÉTAPES POUR CONSERVER DES DOCUMENTS NUMÉRIQUES EXPLOITABLES²

Cette annexe peut être utilisée comme un document à distribuer aux responsables et au personnel des TIC des ministères.

La plupart des organisations créent et stockent des volumes de plus en plus importants d'informations numériques, à savoir d'informations créées à l'aide d'ordinateurs. Ces informations sont souvent d'une importance essentielle en tant que source d'information et de preuve. La perte de ces informations compromettrait gravement la capacité de l'organisation à fonctionner de façon efficace et efficiente, et pourrait générer des dommages critiques pour la réputation de l'organisation auprès de ses clients, citoyens et parties prenantes.

Les informations numériques sont très vulnérables à la perte ou à la destruction délibérée ou accidentelle. L'évolution rapide des technologies peut rendre les anciennes formes d'informations numériques inutilisables et inaccessibles. Il est essentiel que les organisations mettent en place des stratégies permettant de garantir que les informations critiques sont disponibles et exploitables aussi longtemps que nécessaire.

1. Connaître ses documents d'activité

Connaître :

- les systèmes informatiques dont vous disposez et les technologies sur lesquelles ils s'appuient pour fonctionner, car tôt ou tard, ces technologies deviendront obsolètes et feront que les informations stockées dans le système risquent d'être perdues.
- les informations créées et conservées dans chaque système, les formats dans lesquels les informations sont stockées, leur emplacement, la façon dont elles sont décrites et leur raison d'être.
- la façon dont les informations sont liées aux activités de votre organisation, la façon dont elles sont utilisées et la durée pendant laquelle elles doivent être conservées pour répondre aux exigences commerciales et légales.

² Adapté avec l'autorisation du conseil de la State Records Authority of New South Wales, « How to avoid information loss in the digital age » (Comment éviter la perte d'informations à l'ère du numérique), <http://futureproof.records.nsw.gov.au/wp-content/uploads/2010/05/Managing-digital-records-leaflet-Final1.pdf>, et Queensland State Archives Public Records Brief, « Keeping records useable - Ten steps for ensuring the continued accessibility of digital records » (Conserver des documents d'activité exploitables - dix étapes pour garantir l'accessibilité en continu des documents d'activité numériques), http://www.archives.qld.gov.au/publications/PublicRecordsBriefs/Ten_steps_digital_records.pdf (en anglais)

2. Concevoir des systèmes pour prendre en charge vos documents d'activité

De nombreux systèmes informatisés ne sont pas conçus pour conserver un accès aux informations stable sur le long terme. Il est important de reconnaître que l'utilisation des informations sur le long terme est une exigence fonctionnelle à la conception ou à l'achat d'un nouveau système. Dans les cas dans lesquels les données doivent être conservées au-delà de la durée de vie prévue du système (généralement cinq à dix ans), il est essentiel que le système soit en mesure d'exporter les données sous une forme utilisable afin qu'elles puissent être transférées dans de futurs systèmes opérationnels.

3. Limiter le nombre des formats de fichiers que vous utilisez

Les formats de fichiers sont le mécanisme par lequel différents types d'informations numériques sont codés et stockés pour être utilisés. Comme les formats de fichiers deviennent obsolètes avec le temps, il sera nécessaire de faire migrer les données opérationnelles essentielles depuis les anciens formats de fichiers vers des formats nouveaux ou plus stables. Plus votre organisation utilise de formats de fichiers, plus il sera coûteux et laborieux de garantir la longévité de vos actifs informationnels, et plus le risque de perte d'informations vitales sera élevé.

4. Utiliser des formats ouverts

Les formats ouverts ne sont pas détenus par des sociétés de logiciels et sont donc moins vulnérables aux pertes liées à l'obsolescence technologique et découlant de dépendances propriétaires. En règle générale, les formats ouverts sont plus stables et plus faciles à faire migrer. Utilisez dès que possible des formats ouverts tels que PDF, HTML, XHTML, ODF, JPEG et FLAC.

5. Ne pas conserver les informations numériques plus longtemps que nécessaire

Conserver des informations numériques indéfiniment est coûteux et souvent techniquement complexe. Bien que le stockage numérique puisse sembler bon marché, le stockage n'est pas le seul coût associé à la conservation des informations numériques - la gestion et la migration des informations entraînent également des coûts importants. Même avec un stockage bon marché, l'augmentation rapide du volume d'informations numériques signifie que la conservation des données plus longtemps que nécessaire entraînera des coûts insoutenables. Vous devez connaître la durée de conservation des informations et la façon dont elles peuvent être trouvées, et disposer de systèmes et de processus permettant de garantir que les informations importantes sont correctement conservées tandis que les informations sans importance sont mises au rebut en temps voulu et de façon gérée.

6. Savoir où se trouvent toutes vos informations et les garder sous contrôle

Les organisations peuvent stocker leurs informations dans des bases de données centralisées, sur des supports de stockage mobile, sur des bandes de sauvegarde, dans des dossiers personnels ou partagés, ou elles peuvent les faire gérer par des prestataires de services sous contrat. De plus en plus souvent, elles peuvent être

stockées dans ce que l'on appelle le « nuage » Internet, ce qui peut signifier que les informations sont en fait stockées dans un autre pays. Plus les modalités de stockage de vos informations sont diverses et non gérées, plus vos informations risquent d'être perdues.

7. Bien décrire vos informations

Les informations numériques doivent être bien décrites (à l'aide de métadonnées) afin de pouvoir être trouvées, utilisées et gérées. Les bonnes métadonnées comprennent : des titres significatifs et précis ; des informations sur les conditions et les restrictions d'accès et d'utilisation ; des informations qui relient les données à leur contexte opérationnel ; et des informations sur la durée de conservation des données.

8. Réduire les doublons

Les informations numériques peuvent être facilement copiées. Les organisations disposent souvent de milliers de copies d'un même élément d'information, stockées à différents endroits. Gérer et stocker de nombreuses copies des mêmes données est un gaspillage d'argent et peut entraîner une certaine confusion chez le personnel et les clients. Il est impératif de mettre en place des systèmes et des processus permettant d'identifier et d'éliminer les doublons, de contrôler les versions et de garantir l'intégrité et l'authenticité des documents d'activité officiels.

9. Gérer la migration

La migration des données depuis des anciens formats de fichiers et supports de stockage bientôt obsolètes vers des formats nouveaux ou plus stables est une activité indispensable de préservation des informations qui doivent être conservées au-delà de quelques années. La migration est un processus à haut risque qui peut menacer l'intégrité, voire l'existence, d'informations importantes. Les projets de migration doivent par conséquent être soigneusement planifiés et exécutés pour atténuer les risques et protéger l'authenticité, l'intégrité et l'accessibilité des informations.

10. Ne pas attendre qu'il soit trop tard pour garantir la survie de vos informations - ne pas remettre au lendemain ce qu'on peut faire aujourd'hui !

Vous devez dès aujourd'hui garantir la longévité des informations opérationnelles essentielles, sans laisser vos futurs collègues régler rétrospectivement les problèmes que vous leur avez laissés. Réfléchissez systématiquement à la façon dont les systèmes et processus en vigueur peuvent contribuer à garantir que les informations opérationnelles essentielles seront disponibles à l'avenir, lorsqu'elles seront nécessaires. Une bonne gestion de vos informations dès le départ vous permettra d'économiser beaucoup d'argent et d'éviter des problèmes ultérieurs.

ANNEXE D : REFORMATAGE NUMÉRIQUE D'ENREGISTREMENTS AUDIOVISUELS ANALOGIQUES

Tout comme les informations numériques peuvent devenir inaccessibles en raison de l'obsolescence technologique, les enregistrements audio, les films et les bandes vidéo analogiques peuvent l'être aussi. Les bandes audio des bobines et des cassettes, les bandes vidéo VHS et Beta et les films 16 mm et 35 mm sont tous en train de devenir obsolètes. Non seulement les supports de bande physique se détériorent, mais l'industrie ne soutient plus activement les formats correspondants. Il devient presque impossible d'acheter ou de réparer des cassettes audio et des bandes vidéo analogiques et l'équipement nécessaire pour lire ces bandes et cassettes. Toute organisation qui détient du matériel audiovisuel analogique de valeur devra transférer son contenu vers un format numérique afin que ce contenu soit accessible à l'avenir.

L'Association internationale d'archives sonores et audiovisuelles (IASA) a élaboré des directives sur la création et la préservation de copies numériques à partir d'originaux analogiques (consultez la liste de lecture et également l'article de Carl Fleischhauer). Les directives de l'IASA couvrent :

- la sélection et la préparation de la meilleure copie disponible pour la numérisation
- l'optimisation de l'extraction du signal
- les caractéristiques techniques du convertisseur analogique-numérique
- le format audio cible : format PCM Broadcast Wave Format linéaire
- les recommandations de stockage.

L'achat et l'utilisation des équipements nécessaires pour reformater numériquement les enregistrements audiovisuels analogiques aux normes d'archivage peuvent être très coûteux. À moins que votre organisation ne dispose d'une très grande quantité de ce type de matériel, il peut être conseillé d'externaliser le travail de reformatage numérique des collections audiovisuelles analogiques. Dans les îles du Pacifique, cela peut signifier l'externalisation du travail à un prestataire de services extra-territorial. Comme pour tout accord d'externalisation, il convient d'accorder une attention particulière à la réalisation des contrôles préalables, à la négociation des contrats, à la gestion des contrats et à l'assurance qualité des services fournis.

Il est également possible d'acheter des logiciels et des équipements relativement bon marché pour effectuer cette tâche en interne, mais les résultats peuvent être de qualité incertaine et inégale. Consultez la liste de lecture et les articles de Mike Jasper et Alexis Lawrence.

Une fois que les enregistrements audiovisuels ont été convertis sous forme numérique, les copies numériques qui en résultent devront bien sûr être préservées en utilisant les stratégies de préservation numérique décrites dans cette directive.



La Boîte à outils d'archivage pour une bonne gouvernance est produite par la Branche Régionale pour le Pacifique du Conseil International d'Archives, avec l'aide des Archives nationales d'Australie et d'AusAID.