

Annexe 21c

Politiques de maintenance et de conservation des documents d'archives numériques⁵⁸⁸

Préambule

Cette annexe comporte une liste de politiques de conservation largement inspirées des *Directives pour la sauvegarde du patrimoine numérique*⁵⁸⁹ élaborées pour l'UNESCO. Ces dernières offrent un cadre pour la description des politiques et des méthodes de conservation des documents d'archives numériques qui peuvent être utilisées pour protéger et maintenir l'accessibilité des copies authentiques des documents tout au long de la chaîne archivistique.

La liste des politiques adoptées par InterPARES 2 est divisée en deux catégories : a) les politiques de maintenance et b) les politiques de conservation.

A. Politiques de maintenance

Les politiques de maintenance⁵⁹⁰ sont les conditions minimales à satisfaire pour protéger et maintenir l'accessibilité aux copies authentiques de documents d'archives numériques. Il existe huit politiques de maintenance de base. Toutes sont nécessaires pour garantir que les composants des documents d'archives continuent d'exister jusqu'à l'application des politiques de conservation.

A1. Définition des responsabilités

Une personne ou un service doit être désigné comme responsable de la gestion du stockage et de la protection des documents. Il s'agit d'une responsabilité technique qui requiert des compétences spécifiques, des ressources dédiées et un plan idoine. Cette politique peut être conduite en recrutant un collaborateur compétent qui y sera exclusivement dédié ou en assignant à un collaborateur ou un service existant un volume horaire à y consacrer.

A2. Mise en place d'une infrastructure technique adaptée

Cette infrastructure recouvre toutes les ressources matérielles et administratives qui rendent possibles les processus d'archivage et/ou de maintenance des documents d'archives (locaux, matériel informatique, réseaux informatiques et personnel auxiliaire requis pour leur maintenance).

A3. Maintenance, support et remplacement des systèmes

Prévoir la mise en œuvre d'un plan pour maintenir, mettre à niveau et/ou remplacer le matériel et les logiciels.

A4. Migration périodique des données sur de nouveaux supports de stockage

Prévoir la mise en œuvre d'un plan pour copier les données depuis un support de stockage sur un autre afin de

se prémunir contre les effets de la détérioration des supports. Ces migrations doivent être effectuées de manière systématique.

A5. Respect de conditions climatiques appropriées pour les supports de stockage

Le taux de détérioration des supports peut être considérablement réduit en respectant des conditions climatiques adaptées. Par exemple, la poussière, un taux d'humidité ou une température trop élevés mettent en danger les supports.

A6. Redondance et localisation géographique

La duplication des objets numériques et le stockage sur différents supports des copies multiples en résultant, permettent de les protéger de la défaillance des supports. Le stockage dans différents locaux les protège contre de mauvaises conditions ambiantes, l'incendie, les inondations, etc., sur un site donné.

A7. Sécurité des systèmes

Des mesures de contrôle doivent être mises en place afin de garantir que les composants numériques des documents ne soient accessibles qu'aux utilisateurs et/ou processus autorisés. Ces mesures comprendront notamment un accès restreint aux locaux où sont conservés les ordinateurs ainsi qu'aux documents numériques sur les ordinateurs eux-mêmes. Différentes solutions sont envisageables, comme l'utilisation de mots de passe et/ou l'authentification biométrique pour se connecter au système.

A8. Plan de prévention des sinistres et plan de restauration

Les politiques répertoriées ci-dessus sont conçues pour minimiser les pertes accidentelles de données et maximiser la longévité des supports. Toutefois, même dans des conditions de stockage optimales et avec d'excellents protocoles de manipulation, des sinistres peuvent se produire. Un plan de restauration doit comporter des procédures détaillées pour restaurer les systèmes endommagés et pour guider la remise en état du système d'archivage à la suite d'un sinistre.

B. Politiques de conservation

Outre les politiques de maintenance, tous les producteurs de documents doivent mettre en place un système d'archivage de confiance⁵⁹¹ fondé sur une ou plusieurs politiques de conservation⁵⁹². Douze politiques de conservation sont répertoriées ci-dessous en section B, divisées en quatre grandes catégories. En pratique, il est vraisemblable qu'un producteur mettra en œuvre au moins deux politiques de conservation en plus des politiques de maintenance répertoriées en section A.

B1. Utilisation de normes et de standards

L'utilisation de standards et/ou normes largement disponibles et maintenus augmente la probabilité de stabilité et de support technique à long terme. Il pourra s'agir de normes⁵⁹³, si elles ont été officiellement agréées, ou de standards⁵⁹⁴ s'ils ont été adoptés par un groupe d'utilisateurs. Les normes et standards peuvent concerner de multiples facettes d'un système de conservation : méthodes d'encodage, formats de fichiers, supports de stockage, etc. Le respect de normes ou de standards est également susceptible de simplifier la mise en place des politiques de conservation ultérieures et/ou d'en maximiser l'efficacité. La normalisation peut être conduite de manière *prospective*, en limitant les formats dans lesquels les documents numériques peuvent être versés au service d'archives, ou de manière *retrospective*, en convertissant dans des formats standards les fichiers reçus dans d'autres formats.

B1.1. Formats auto-descriptifs (conservation dans un format pérenne, balisage)

Cette méthode consiste à analyser et baliser les documents d'archives afin de permettre la description des fonctions, des relations et de la structure d'éléments particuliers. La représentation du contenu peut ainsi s'affranchir d'applications logicielles spécifiques et être réalisée en utilisant différentes applications au fil des évolutions technologiques.

B1.2. Encapsulation

Méthode utilisée pour lier un document et les moyens d'y accéder, normalement dans un *conteneur* qui décrit le document d'une manière qui peut être comprise par un large éventail de technologies (comme un document XML). Le conteneur inclut souvent des métadonnées qui décrivent les outils à utiliser ou y renvoient par un lien.

B1.3. Limiter l'éventail de formats à gérer (normalisation)

Stocker des documents dans un nombre limité de formats seulement⁵⁹⁵. La sélection des formats acceptables peut être étendue en vue d'y inclure de nouveaux formats propriétaires ou de nouvelles générations de formats propriétaires existants, ou être limitée à des formats non propriétaires, afin de renforcer la normalisation. Un exemple de cette approche est l'*encodage pérenne*, qui préconise d'encoder les documents en se conformant à des traitements de données bien connus, normes d'encodage de caractères comme ASCII ou UTF-8 ou XML pour les objets, notamment.

B1.4. Conversion

Transférer des documents numériques d'une génération de matériel ou de logiciels à une autre. À distinguer du *rafraîchissement* qui consiste à copier le flux de données d'un support numérique sur un autre. La conversion consiste à transformer le format logique d'un objet numérique de façon à ce que l'objet conceptuel puisse être correctement restitué ou affiché par le nouveau matériel ou le nouveau logiciel. La méthode de conversion la plus couramment proposée consiste à transformer de manière permanente un format logique en un autre en fonction de l'évolution technologique, pour que tous les objets convertis puissent être restitués à l'aide de la technologie en place. Il est également possible de proposer un modèle de « conversion à la demande » ou de « conversion au point d'accès » (voir la section B2.4 ci-dessous).

B2. Dépendance technologique

Ce type de méthodes de conservation continue de reposer sur les matériels et/ou logiciels initiaux sans modifier les documents.

B2.1. Préservation de la technologie

Maintenir le matériel ou le logiciel d'origine, avec lesquels les documents numériques ont été représentés.

B2.2. Dépendance à l'égard de la rétrocompatibilité

Faire confiance à la capacité d'un logiciel d'interpréter et de présenter correctement les composants numériques de documents créés avec des versions antérieures de celui-ci. Avec cette méthode, la présentation se limite à une conversion temporaire à des fins de visualisation, alors que la conversion transforme de manière permanente les documents dans le format supporté par la version actuelle du logiciel.

B2.3. Ré-ingénierie logicielle

Modifier les logiciels au fil des évolutions technologiques. À ce titre, cette stratégie est similaire à la transformation des formats de documents, abordée en sections B1.4 et B2.2. Elle peut notamment comprendre la nouvelle compilation du code source pour une nouvelle plateforme ou encore un nouveau codage du logiciel dans un autre langage de programmation.

B2.4. Visualiseurs et conversion au point d'accès

Recours à des outils logiciels ou à des méthodes de transformation qui fournissent une accessibilité temporaire au document, en utilisant le flux de données initial.

B2.5. Émulation

Utilisation de logiciels permettant à une technologie de se comporter comme une autre. En d'autres termes, faire en sorte que les technologies futures se comportent comme l'environnement originel du document d'archives numérique, de sorte que le document original puisse être présenté/visualisé dans sa manifestation originelle à partir des flux de données initiaux ou convertis.

B3. Méthodes non numériques

Copier les documents d'archives numériques sur des supports analogiques relativement stables, comme le papier ou le microfilm ; transférer la charge de la conservation à une copie analogique en lieu et place de l'objet numérique. Cette méthode entraîne la perte de toutes les fonctionnalités offertes par le logiciel, ainsi de la possibilité de manipuler le document.

B4. Restauration des données (archéologie numérique)

Récupérer des documents sous la forme de bits à partir des supports, puis suivre une série d'étapes pour restaurer l'intelligibilité des documents récupérés. Cette méthode est le plus souvent utilisée pour récupérer les données sur des supports détériorés, endommagés ou défectueux, mais elle peut également être utilisée pour sauver des documents dans des formats obsolètes.

⁵⁸⁸ Adapté de : Kevin Glick, « Electronic Records Preservation Strategies », (rapport non publié, 2006).

⁵⁸⁹ Colin Webb (2003), *Directives pour la sauvegarde du patrimoine numérique*. Document élaboré par la Bibliothèque nationale d'Australie pour la Division de la société de l'information, Organisation des Nations Unies pour l'éducation, la science et la culture, rapport n° CI-2003/WS/3. Le document est disponible en français à l'adresse suivante : <http://unesdoc.unesco.org/images/0013/001300/130071f.pdf>.

⁵⁹⁰ Une politique de maintenance est « un ensemble cohérent d'objectifs et de méthodes pour protéger et maintenir l'accessibilité des copies authentiques de documents d'archives numériques au cours des premières phases de la chaîne archivistique. » (glossaire français, op. cit.).

⁵⁹¹ Un système d'archivage de confiance est « L'ensemble des règles qui contrôlent la conservation et l'utilisation des documents d'archives du producteur et fournissent une présomption d'authenticité de ces documents, ainsi que les outils et mécanismes utilisés pour appliquer ces règles. » (*Ibid.*)

⁵⁹² Une politique de conservation est « Un ensemble cohérent d'objectifs et de méthodes pour protéger et maintenir (c'est-à-dire, préserver l'authenticité et garantir l'accessibilité) dans le temps des composants numériques et informations associées des archives définitives, et pour reproduire les documents et/ou les ensembles organiques de documents authentiques. » (*Ibid.*)

⁵⁹³ Une norme est « Un standard publié par un organisme de normalisation, national (par exemple, ANSI), multinational (par exemple, CEN) ou international (par exemple, ISO). » (*Ibid.*) Deux normes récentes de formats de fichiers sont le

PDF/A (norme PDF pour l'archivage) et l'ODF (format OASIS OpenDocument).

⁵⁹⁴ Un standard est « Un ensemble de recommandations qui n'est publié par aucun organisme de normalisation mais est néanmoins largement utilisé et considéré comme une norme par ses utilisateurs. » Citons par exemple les formats PDF, TIFF, DOC et ZIP.

⁵⁹⁵ Pour une analyse détaillée des enjeux et évolutions actuels du choix de formats de fichier, de conteneurs, de balisage et d'encodage, et des recommandations pour l'élaboration et la mise en place de politiques pour la sélection de formats de fichiers numériques pour la conservation sur le long terme, voir : Evelyn Peters McLellan (2006), « InterPARES 2 Project – General Study 11 Final Report : Selecting Digital Formats for Long-Term Preservation », http://www.interpares.org/display_file.cfm?doc=ip2_gs11_final_report_english.pdf. Une version française est également disponible : http://www.interpares.org/display_file.cfm?doc=ip2_gs11_final_report_french.pdf.

Annexe 21b Conditions de base
requis...

InterPARES 2
Annexe 21c Politiques de maintenance...

Glossaire