Manuel de préservation numérique

Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Bienvenue dans la deuxième édition révisée du manuel de préservation numérique. Il s'agit d'une base de connaissances essentielle pour la préservation numérique, évaluée par des pairs et librement accessible à tous.

« Dans l'ensemble, les améliorations apportées au manuel en font, à mon avis, l'un des outils les plus utiles et les plus souples pour identifier, comprendre et maîtriser les approches pratiques des différents défis de la préservation numérique. Il utilise un langage accessible, une terminologie claire et fournit des liens utiles vers des études de cas et des lectures complémentaires qui seront utiles aux étudiants comme aux professionnels ». *Stefanie Davidson, West Yorkshire Archive Service*

L'information numérique est de plus en plus importante pour notre culture, notre base de connaissances et notre économie. Ce manuel, compilé pour la première fois par Neil Beagrie et Maggie Jones en 2001, est maintenu et mis à jour par la DPC. Sa révision complète (2ème édition) a élargi et mis à jour son contenu pour couvrir plus de 30 sections principales (voir Table des matières). Cette 2ème édition a été compilée avec la contribution de 45 professionnels et experts en préservation numérique sous la direction de Neil Beagrie en tant que rédacteur en chef et de William Kilbride en tant que président du conseil d'administration et du conseil consultatif. Le manuel est un guide pratique qui fait autorité au niveau international sur la gestion des ressources numériques dans le temps et sur les questions liées à la pérennité de l'accès à ces ressources Il intéressera tous ceux qui participent à la création et à la gestion des objets numériques.

Informations sur la traduction française

La traduction française du manuel de préservation numérique a été réalisée dans le cadre des travaux de la Cellule nationale de veille sur les formats entre juillet 2020 et janvier 2021. Lancée en 2019, la Cellule nationale de veille sur les formats (CNVF), sous l'égide de l'association française Aristote et de son groupe de travail sur la Pérennisation de l'information numérique (PIN), regroupe à ce jour une douzaine de partenaires. Ses principaux objectifs sont la mutualisation des activités de veille sur les formats, la sensibilisation des professionnels sur le sujet, la contribution ou l'influence sur les outils associés. Elle ambitionne d'être un interlocuteur francophone reconnu dans les travaux internationaux sur ces sujets.

Le manuel a été traduit en français par : Marion Humbert (Archives départementales de la Moselle), Anne Jolly (ministère des Armées), Thomas Ledoux (Bibliothèque nationale de France), Wilfried Prieur (ministère des Armées), Stéphanie Roussel (mintika), Edouard Vasseur (École nationale des Chartes).

Une première traduction automatique a été réalisée avec le logiciel DeepL, puis a été revue et corrigée par le groupe de travail.
Parmi les choix de traduction effectués, on signalera tout particulièrement les points suivants :

- Le terme de **préservation numérique** a été retenu pour traduire *digital preservation*, en cohérence avec les autres travaux de traduction de la Cellule nationale de veille sur les formats ;
- Le terme d'**entrepôt numérique** a été retenu pour traduire *repository*, de préférence à Système d'archivage électronique (SAE) qui est propre à la sphère archivistique ;
- Le terme **Archive** au sens de la norme OAIS a été traduit par le terme Archive (avec une majuscule) ou par entrepôt numérique en fonction du contexte ;

- L'expression **objet numérique** a été retenue pour traduire *digital material* ;
- L'expression *dark archive* a été conservée en langue anglaise.

S'agissant d'un manuel anglo-saxon, certaines parties du manuel, comme celles relatives au contexte juridique et au contexte normatif, sont peu adaptées au contexte français. Pour des ressources plus spécifiques à ce dernier, on se reportera aux travaux de la Cellule nationale de veille sur les formats disponibles en ligne.

Lors de la traduction, il a parfois été nécessaire d'ajouter des éléments complémentaires propres au contexte français ou d'expliciter certains éléments. Ces ajouts, propres à la version française, ont été signalés soit par un astérisque soit par la mention NdT (note du traducteur).
Enfin, on signalera que le développement et l'explicitation des sigles et acronymes ont été systématiquement renvoyés au glossaire. Les entrées en langue anglaise ont également été conservées dans le glossaire.

Table des matières

Manuel de préservation numérique 1
Informations sur la traduction française 2
Table des matières 4
Introduction 6
Comment utiliser le manuel ? 9
Élaboration et remerciements 13
Récapitulatif sur la préservation numérique 19
Pourquoi la préservation numérique est-elle importante ? 20
Problèmes de préservation 24
Pour commencer 36
Stratégies institutionnelles 45
Politiques et stratégies institutionnelles 46
Collaboration 53
Sensibilisation 59
Achats et recours à des tiers 64
Audit et certification 74
Respect de la législation 84
Gestion des risques et du changement 93
Formation et développement des compétences du personnel 99
Normes et bonnes pratiques 106
Analyse de rentabilité, avantages, coûts et impact 115
Activités organisationnelles 126
Création d'objets numériques 127
Acquisition et évaluation 136
Préservation et (ré)évaluation 158
Stockage 162
Anciens supports 172
Planification de la préservation 176
Mesures de préservation 181
Accès 187
Métadonnées et documentation 193
Solutions et outils techniques 200
Outils 201
Intégrité et empreintes 207
Formats de fichiers et normes 212

Sécurité de l'information 220
Services *cloud* 226
Criminalistique numérique 233
Identifiants pérennes 238
Préservation du contenu spécifique 243
Périodiques électroniques 244
Images animées et son 251
Archivage web 261
Glossaire 272

Des mises à jour et des ajouts seront faits de manière régulière. Toute nouvelle section prévue dans les prochaines publications de la DPC sera indiquée comme « à venir ».

Introduction

Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Le volume d'information existant sous forme numérique est en augmentation rapide.

Qu'elles soient créées à la suite de la numérisation de collections non nativement numériques, sous forme de publication ou d'œuvre d'art numérique, ou dans le cadre des activités quotidiennes d'une organisation ou d'un individu, de plus en plus d'informations sont créées sous forme numérique et le rythme auquel elles sont créées s'accélère. On observe en parallèle l'émergence d'une conscience des défis majeurs posés par la garantie d'un accès continu à ces objets, même à court-terme.

La combinaison de ces deux facteurs est à la fois un défi et une inquiétude. D'une part, la technologie numérique offre des possibilités considérables pour fournir un accès rapide et efficace à l'information. D'autre part, il existe une menace très réelle que les objets numériques soient créés d'une manière telle que non seulement leur viabilité à court terme ne puisse être assurée, mais qu'il n'y ait en outre aucune perspective pour que les générations futures puissent y avoir également accès.

Besoin de conseils

Toutes les organisations de tous les secteurs ont créé et créent des documents numériques. Ceux-ci peuvent être créés en tant que documents d'activité de l'organisation, ils peuvent être créés en numérisant des collections non numériques afin d'en améliorer l'accès, ou ils peuvent être créés numériquement (« nativement numériques »). Quoi qu'il en soit, dès qu'ils voient le jour, ces documents doivent être gérés le plus tôt possible au cours de leur cycle de vie, de préférence au stade de leur conception, et, si cela n'est pas possible, ils devront rester accessibles aussi longtemps que nécessaire. L'expérience pratique et l'expertise dans ce domaine sont encore relativement limitées, si bien qu'il existe un besoin évident d'orientations pour s'assurer que des opportunités significatives ne sont pas contrecarrées par des menaces tout aussi importantes.

Ce manuel vise à identifier les bonnes pratiques en matière de création, de gestion et de préservation des objets numériques et à fournir une série d'outils pratiques pour faciliter ce processus. Bien qu'il reste des défis à relever, nous pouvons citer de nombreux exemples de bonnes pratiques et donner des suggestions pour permettre aux institutions de s'initier à la préservation numérique. En fournissant un aperçu stratégique des questions clés, des discussions

et des conseils sur les stratégies et les actions à mener, ainsi que des indications sur les principaux projets et rapports, le manuel vise à fournir des conseils aux institutions et aux individus ainsi qu'une gamme d'outils pour les aider à identifier et à prendre les mesures appropriées.

Public et objectifs

La préservation numérique présente de nombreux parallèles avec la préservation traditionnelle en ce qui concerne les principes généraux, mais elle diffère nettement au niveau opérationnel et tout particulièrement dans le large éventail de décideurs qui jouent un rôle crucial aux différents stades du cycle de vie d'une ressource numérique. C'est pourquoi ce manuel s'adresse à un très large public. En premier lieu, il vise à fournir des conseils aux institutions de niveau international, national, régional ou local, impliquées dans la création et/ou l'acquisition d'objets numériques ou qui envisagent de le faire. Au sein de ces institutions, le manuel s'adresse à la fois aux administrateurs et aux professionnels et il est donc structuré de manière à mêler des vues d'ensemble stratégiques de haut niveau et des orientations détaillées. En outre, le manuel s'adresse aux prestataires de services qui peuvent être en mesure de fournir tout ou partie des services nécessaires à la préservation des objets numériques. Il s'adresse également aux organismes de financement qui devront être conscients de ce qu'implique la création d'objets numériques. Enfin, il intéressera les créateurs de données dont l'implication dans la préservation de leurs objets numériques reste cruciale, bien qu'elle soit limitée par les moyens généraux de leur organisation.

Le manuel prend pleinement en compte le fait que ces groupes peuvent avoir des intérêts et une implication variés par rapport aux objets numériques et au cours du temps. En adoptant une approche par le cycle de vie en matière de préservation numérique, il veut aider à identifier les dépendances, les obstacles et les mécanismes, afin de faciliter la communication et la collaboration entre les acteurs.

Le manuel doit être adapté aux besoins individuels et doit servir de catalyseur aux actions des institutions en leur sein et entre elles, y compris entre les institutions où la préservation numérique peut être externalisée ou envisagée uniquement à court terme.

Les questions majeures en lien avec la préservation numérique sont naturellement universelles ; aussi les exemples de bonnes pratiques, les travaux de recherche et les conseils ou orientations produits à l'échelle du globe ont été mobilisés. Toutefois, en ce qui concerne le contexte, l'étude se concentre sur le Royaume-Uni et certains exemples, comme la législation, sont spécifiques au Royaume-Uni. Le texte du manuel indiquera l'accent mis sur le Royaume-Uni chaque fois que nécessaire. Nous espérons néanmoins que le manuel sera pertinent pour un public international, car de nombreux modèles et références fournis ne proviennent pas du Royaume-Uni et sont de toute façon applicables à n'importe quel pays. Quel que soit leur pays d'origine, les utilisateurs du manuel devront l'adapter à leurs besoins spécifiques.

Bien que les questions soient complexes et qu'il reste beaucoup à clarifier (certains points ne seront peut-être jamais définitivement résolus), le propos général du manuel vise à montrer que beaucoup de choses ont déjà été réalisées et que beaucoup d'autres peuvent être entreprises immédiatement par tous ceux qui sont impliqués dans la création et/ou l'acquisition d'objets numériques. Ces actions contribueront à protéger l'investissement initial que représente la création d'objets numériques et offriront des perspectives considérablement améliorées à long terme.

Principes directeurs

Les principes suivants ont guidé l'élaboration du manuel:

- Être en cohérence avec la mission de la DPC en étant indépendant des fournisseurs, en encourageant la participation active de la communauté et en partageant les connaissances et les meilleures pratiques.

- Répondre aux besoins de la communauté de la préservation numérique et lui être utile. La communauté des utilisateurs actuels et potentiels du manuel sera interrogée et les commentaires seront utilisés pour améliorer son développement. De futures évaluations enrichiront les retours d'expérience et les analyses d'impact.

- Être informé, à jour, concis et équilibré. Les chapitres fourniront des aperçus concis des sujets avec une sélection de lectures complémentaires et d'études de cas offrant la possibilité d'approfondir davantage les sujets. La sélection de lectures complémentaires et d'études de cas sera guidée par l'utilité, l'actualité et la pertinence de leur contenu.

- Réduire les obstacles de la participation à la préservation numérique par une rédaction dans un style informatif mais accessible à un large public.

- Proposer un contenu évalué par des pairs, faisant autorité, avec des critères de qualité explicites, tels que définis dans les *Notes pour les auteurs, relecteurs et éditeurs de la DPC* et dans le plan de projet du manuel.

- Travailler ensemble et créer des synergies avec les rapports de la série *Tech Watch* de la DPC[1]. Les rapports fournissent des approfondissements dans des domaines spécifiques de la préservation de certains contenus (par exemple, le courrier électronique) ou sujets (par exemple, la criminalistique numérique) qui peuvent être cités ou servir de source pour des études de cas dans le manuel.

- Promouvoir d'autres ressources (par exemple, le registre COPTR (*Community Owned digital Preservation Tool Registry*) qui peuvent être des moyens à privilégier pour se renseigner sur des domaines particuliers en dehors du manuel.

- Promouvoir d'autres ressources (par exemple, des rapports, des conseils) qui sont accessibles sur le web et gratuites pour les utilisateurs, afin que, dans la mesure du possible, tout le contenu externe sélectionné soit aussi accessible que le manuel lui-même.

- Être développé par la DPC dans un souci de facilité de maintenance, d'économie et de durabilité à long terme.

[1]Les rapports *Tech Watch* de la DPC fournissent des conseils faisant autorité sur une variété de sujets liés à la préservation numérique choisis par les membres de la DPC. Ils sont conçus pour informer les membres de la DPC, et ceux de la communauté au sens large, d'une manière accessible et facile à assimiler.

Développement futur et soutien

À l'issue de la deuxième édition, la DPC a l'intention d'intégrer sa maintenance et sa révision à sa série de rapports *Tech Watch* sur la préservation numérique. Un comité consultatif de rédaction de DPC sera chargé de superviser la révision des chapitres du manuel et des rapports ou de commander de nouveaux sujets pour les rapports ou le contenu du manuel.

Comment utiliser le manuel ?

Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Mise en pratique

Les besoins des institutions concernant les objets numériques qu'elles créent et acquièrent varient considérablement. Ce manuel est destiné à faire le lien entre des aperçus généraux de haut niveau et des lignes directrices explicites et détaillées applicables aux besoins d'une institution spécifique. Il est souhaitable que les aperçus stratégiques soient liés aux activités opérationnelles afin de renforcer la nécessité d'élaborer des procédures pratiques solidement ancrées dans la mission de l'institution. Le manuel fournit des indications sur les endroits où trouver des conseils supplémentaires et des aides pour développer les politiques et les pratiques qui sont les plus applicables à une institution en particulier.

Idéalement, le manuel devrait être utilisé pour aider à amorcer les réflexions, à améliorer la compréhension générale, à promouvoir la formation et à agir comme un catalyseur pour de nouvelles actions. Rien n'exclut que chaque organisation doive finalement engager les ressources nécessaires à un plan d'action, mais ce manuel est destiné à huiler les rouages de ce processus.

Ressources du manuel

Le manuel préconise une approche globale de la préservation des ressources numériques fondée sur des principes et des politiques solides plutôt que sur des formules prescriptives. L'importance cruciale de la préservation numérique étant de plus en plus largement reconnue, des ressources précieuses sont publiées en grand nombre et à un rythme rapide. Bien qu'elles soient potentiellement très utiles, leur prolifération même peut déconcerter et il est difficile de définir lesquelles sont susceptibles d'être les plus applicables dans une situation donnée.

En sélectionnant des ressources et des études de cas clés dans chaque section, le manuel devrait faciliter la navigation dans les sources existantes en matière de conseils, d'orientations et d'options. En plus de pointer vers les sources existantes, nous avons utilisé une combinaison d'arbres de décision, de listes de contrôle sommaires, d'exemples et d'études de cas sélectionnés ainsi que de commentaires. Ces éléments sont destinés à stimuler et à favoriser la réflexion et la discussion, mais surtout à inciter les institutions à prendre des mesures pour élaborer des politiques et des stratégies de gestion de la préservation numérique adaptées à leurs besoins.

Les ressources ont été regroupées en catégories et signalées par les icônes suivantes :

Publications	**Outils**	**Ressources web**	**Vidéos et webinaires**	**Études de cas**

Aperçu général

Le manuel s'adresse à un public large et diversifié, depuis ceux qui commencent seulement à envisager de gérer des documents numériques jusqu'aux professionnels qui ont déjà accumulé une expérience théorique et/ou pratique considérable. Il a été rédigé dans l'intention de permettre un accès rapide et facile aux sections appropriées.

Chaque section est précédée d'une présentation rapide sur le public principal visé, son niveau de connaissance présumé et l'objectif de la section. Le tableau ci-dessous vous aidera à décider quelles sections sont susceptibles d'être les plus pertinentes pour vous. Il ne se veut toutefois pas rigide et toute personne qui le souhaite peut bien sûr lire le manuel dans son intégralité !

Les profils d'administrateurs généraux, d'encadrants et de personnels opérationnels, issus du cadre de compétences DigCurV pour la préservation numérique (DigCurV, 2013), ont été utilisés dans le cadre plus large du manuel comme des classifications d'audience primaire.

Tous les lecteurs sont encouragés à lire la section Introduction.

La section Glossaire fournit des explications concises des concepts clés et des définitions des acronymes et des abréviations utilisés par les organisations et les projets tout au long du manuel.

Sections recommandées par type de public

Public	**Chapitres recommandés**
Toute personne ayant besoin d'une introduction au sujet	Récapitulatif sur la préservation numérique Pour commencer
Créateurs et éditeurs	Récapitulatif sur la préservation numérique Activités organisationnelles Solutions et outils techniques
Financeurs	Récapitulatif sur la préservation numérique
Encadrants (DigCurV : *Manager Lens*)	Stratégies institutionnelles Activités organisationnelles

	Solutions et outils techniques
Personnels opérationnels (DigCurV : *Practitioner Lens*)	Pour commencer Activités organisationnelles Solutions et outils techniques Préservation du contenu spécifique
Administrateurs généraux (DigCurV : *Executive Lens*)	Stratégies institutionnelles
Prestataires de services	Stratégies institutionnelles Activités organisationnelles Solutions et outils techniques

Réutilisation du manuel

Réutilisation en anglais

Le texte du manuel est disponible en anglais sous une licence *Open Government Licence* v3.0 http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/ afin qu'il puisse être réutilisé aussi largement que possible. Vous pouvez incorporer les sections du manuel dans des manuels de formation et d'autres documents. Veuillez utiliser cette formule de remerciement pour la réutilisation : *Digital Preservation Handbook*, 2e édition, https://www.dpconline.org/handbook *Digital Preservation Coalition* © 2015 sous licence *Open Government Licence* v3.0[2].

digitalbevaring.dk a aimablement donné l'autorisation d'utiliser les illustrations des titres de section du manuel et les icônes des ressources et études de cas, produites par Jørgen Stamp. Ces ressources sont protégées par le droit d'auteur de https://digitalbevaring.dk/ et partagées sous une licence CC BY 2.5 Danemark (illustrations) https://creativecommons.org/licenses/by/2.5/dk/deed.en_GB et une licence CC0 1.0 (icônes) https://creativecommons.org/share-your-work/public-domain/cc0. Elles peuvent également être réutilisées si la ligne de crédit et la mention utilisées dans le manuel sont incluses.

Toutefois, veuillez noter que le manuel continuera d'être mis à jour. Aidez les utilisateurs en :

1. Fournissant un lien url vers la dernière édition en ligne du manuel sur le site web de la DPC afin que les utilisateurs puissent vérifier les dernières mises à jour ;
2. Vous inscrivant à la liste de préservation numérique sur Jiscmail afin de recevoir les notifications des dernières modifications et mises à jour du manuel. Pour vous inscrire à la liste, rendez-vous sur le site https://www.jiscmail.ac.uk/cgi-bin/webadmin?A0=DIGITAL-PRESERVATION
3. Citant l'utilisation du manuel.

En plus d'être une bonne pratique, il sera utile à la DPC pour établir son rapport aux financeurs et sponsors du manuel et montrer son impact.

Traduction et réutilisation dans d'autres langues

Les demandes de traduction du manuel dans d'autres langues sont les bienvenues. La DPC suivra les principes et les accords conclus pour la traduction de ses autres publications et conclura un accord officiel pour la traduction afin que toutes les modalités en soient claires à long terme.

[2] NdT : équivalent à la Licence Ouverte v2 (disponible à l'adresse suivante : https://www.etalab.gouv.fr/licence-ouverte-open-licence)

En règle générale, nous souhaiterions :

- Permettre le téléchargement de la traduction sur le site web de la DPC. L'utilisation de nos publications est l'une de nos principales mesures d'impact. Il est donc nécessaire de s'assurer que les statistiques web qui en résultent sont disponibles.

- Attribuer un identifiant numérique (DOI) à la traduction. Cela signifie que les utilisateurs n'ont pas besoin de savoir que le rapport a été diffusé sur le site web de la DPC : d'autres organisations peuvent y donner accès par un lien à partir de n'importe quelle page source. Cela signifie également que nous nous engagerions à maintenir l'url indéfiniment (ce qui signifie que le traducteur n'a pas besoin de le faire).

- Conserver l'image de marque et les crédits existants dans la publication. Nous proposerions un avant-propos ou une introduction spéciale qui pourrait inclure une description de la traduction et nous serions heureux que le logo et la marque soient mis en évidence sur la première page et dans l'ensemble de la publication.

- Conserver un droit de regard sur le résultat final, car la DPC a également un intérêt à veiller à la qualité et à la réputation de la version traduite du manuel. Cela peut inclure l'envoi à un lecteur natif pour une révision indépendante par des pairs.

Références

DigCurV, 2013. *A Curriculum Framework for Digital Curation*.
Disponible à l'adresse suivante : https://www.digcurv.gla.ac.uk/

Élaboration et remerciements

Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Élaboration du manuel

Le manuel a été créé en 2001 sous la forme d'une publication imprimée intitulée *Preservation Management of Digital Materials : a Handbook*, imprimée par *British Library Publishing* (Jones et Beagrie, 2001). Un CD à édition limitée du manuel a été produit en février 2002 pour le lancement de la DPC à la Chambre des communes. La version numérique en ligne du manuel a été mise à disposition sur le site web de la DPC peu de temps après, en mai 2002. Une version archivée de cette « première édition en ligne » du manuel est disponible dans les *UK Web Archive* (UKWA, 2009).

Depuis plus de dix ans, l'édition en ligne du manuel constitue un guide pratique qui fait autorité au niveau international sur la gestion des ressources numériques dans le temps et sur les questions liées à la pérennité de l'accès à ces ressources. Cependant, seuls quelques petits ajouts ont été apportés au manuel depuis sa mise en ligne. La DPC s'est fixé comme objectif stratégique d'obtenir des subventions de recherche pour mettre à jour complètement le manuel et l'intégrer plus étroitement à sa série *Tech Watch* afin de faciliter son actualisation et sa mise à jour.

En 2014, les Archives nationales du Royaume-Uni, en tant que principal financeur, avec une participation supplémentaire de l'*Archives and Records Association*, de la *British Library*, du Jisc et des Archives nationales d'Écosse, ont accordé une subvention de recherche pour commencer la révision complète du manuel et promouvoir son utilisation et son adoption. Ce travail a été entrepris sous la direction de William Kilbride, directeur de la DPC, et de Neil Beagrie, rédacteur en chef, et a été soutenu par un vaste conseil consultatif et un ensemble de contributeurs.

Digital Preservation Coalition est une organisation à but non lucratif qui met gratuitement à disposition le manuel de préservation numérique, *Digital Preservation Handbook,* comme ressource en ligne : celui-ci est très utilisé pour le développement professionnel continu, par les étudiants universitaires et pour la formation à la préservation numérique. Une étude des formations à la préservation numérique par le projet APARSEN a mis en évidence la pauvreté des sujets couverts par la plupart des initiatives (APARSEN, 2012). Il souligne les lacunes actuelles dans

la couverture des sujets et suggère qu'une édition révisée et mise à jour du manuel répondrait aux besoins actuels.

Dans le cadre de la préparation de la 2e édition, une harmonisation a été entreprise entre le manuel et les recommandations d'APARSEN ainsi qu'avec le Programme d'études DigCurV pour la préservation numérique (DigCurV, 2013). Une enquête de satisfaction approfondie a permis d'éclairer les propositions d'amélioration du manuel. Cette étude a bénéficié de la contribution de 285 membres de la communauté via un sondage en ligne (DPC, 2014a) et une consultation publique sur une ébauche de contenu (DPC, 2014b).

Le système de Harvard a été utilisé pour les rubriques « références », dans la continuité de ce qui est fait pour les Rapports de veille technologique de la DPC.

Cet accent mis sur l'engagement de la communauté de la préservation numérique à s'assurer que le manuel représente les besoins d'un large éventail d'organisations et de secteurs s'est poursuivi lors de la révision grâce au recours à des « book-sprints[3] » collaboratifs et au travers du conseil consultatif.

Remerciements

La deuxième édition du manuel de préservation numérique, *Digital Preservation Handbook,* est le fruit d'un projet de recherche commun basé sur un accord de consortium entre les Archives nationales du Royaume-Uni, *The Archives and Records Association*, la *British Library*, Charles Beagrie Ltd, la *Digital Preservation Coalition* et Jisc.

- Le financement du projet a été initié par les **Archives nationales du Royaume-Uni**. Conscientes du défi de la préservation numérique auquel est confronté le secteur des archives et sensibles à leur rôle de leader stratégique du secteur, elles ont offert 50 % des fonds estimés nécessaires au projet (c'est-à-dire le *Gold Sponsorship*), étant entendu que la DPC devrait soit trouver les fonds restants en attirant des sponsors supplémentaires, soit prendre en charge les coûts directement.

- Le conseil d'administration de la **DPC**, désireux de réviser le manuel, a accepté l'offre des Archives nationales du Royaume-Uni, offrant au personnel de la DPC du temps et garantissant que les coûts matériels (tels que le site web) seraient fournis gratuitement. La DPC a également apporté une contribution en espèces provenant des fonds des membres.

- **Charles Beagrie Ltd** a rejoint le consortium en tant que partenaire éditorial pour entreprendre des recherches originales, développer le contenu et la conception, et veiller à la qualité de l'ensemble. Charles Beagrie Ltd a fourni une contribution en nature sous forme de temps de personnel (*Bronze Sponsorship*) pour aider à la finalisation des tâches éditoriales et achever le manuel.

- **Jisc** a rejoint la collaboration très tôt, reconnaissant que les archivistes et les gestionnaires de données de recherche dans l'enseignement supérieur seraient bien placés pour utiliser

[3]Un booksprint est une session de travail intensif organisée pour l'écriture complète d'un livre, par plusieurs auteurs, réunis ou à distance, souvent encadrés par un facilitateur non rédacteur (définition Wikipédia).

le manuel. Les fonds de Jisc (*Silver Sponsorship*) ont permis de développer les trois sections de préservation des contenus spécifiques et la méthodologie du « book-sprint ».

- La ***British Library*** s'est également jointe à la collaboration très tôt, ce qui reflète son rôle dans le manuel original. Les fonds de la *British Library* (*Silver Sponsorship*) ont permis de réaliser une évaluation et une enquête complètes sur les besoins de révision qui ont servi de base à la table des matières révisée.

- Après une période de développement du contenu et de prototypage, l'***Archives and Records Association*** **(ARA)** a rejoint la collaboration en offrant des fonds (*Silver Sponsorship*) provenant de ses propres subventions de recherche, sous la garantie que les projets de textes seraient systématiquement et complètement révisés par des pairs avant leur publication, et que la DPC serait en mesure de mettre le manuel "sur les rails" lors d'une série d'événements de formation accessibles aux membres de l'ARA.

- Enfin, des fonds ont été fournis par les **Archives nationales d'Écosse** (*Bronze Sponsorship*) pour permettre une large diffusion et un lancement rapide.

Le conseil d'administration de la DPC remercie officiellement les partenaires, en notant également les nombreuses contributions supplémentaires, les présentations et les encouragements qu'ils ont apportés pour mener à bien le projet.

Financeurs

Gold sponsor

Silver sponsors

Bronze sponsors

digitalbevaring.dk a aimablement donné l'autorisation d'utiliser les illustrations des titres de section du manuel et les icônes des ressources et des études de cas, produites par Jørgen Stamp. Il s'agit du copyright de digitalbevaring.dk, partagé sous une licence CC BY 2.5 Danemark

(illustrations) https://creativecommons.org/licenses/by/2.5/dk/deed.en_GB et une licence CC0 1.0 (icônes) https://creativecommons.org/share-your-work/public-domain/cc0.

Rédacteur en chef du manuel : Sharon McMeekin (à partir de juin 2016)

Ancien directeur de la rédaction : Neil Beagrie (jusqu'en juin 2016)

Contributeurs du manuel

Matthew Addis	Arkivum Ltd (book sprint 1)
Neil Beagrie	Charles Beagrie Ltd (book sprint 1, book sprint 2, book sprint 3, book sprint 4, Cloud services, Glossary, Handbook 1st edition, Introduction, e-journals preservation case study, Resources and case studies)
Daphne Charles	Charles Beagrie Ltd (Handbook web content design and input, desk research and audience research for Handbook)
Andrew Charlesworth	University of Bristol (Cloud services)
Tracey Clarke	University of Sheffield (book sprint 3)
Glenn Cumiskey	British Museum (book sprint 2)
Stefanie Davidson	West Yorkshire Archive Service (book sprint 1)
Michael Day	British Library (book sprint 1)
Matt Faber	Jisc (book sprint 1, book sprint 2)
Chris Fryer	Parliamentary Archives (book sprint 1)
Dave Govier	Manchester Met Archives Network (book sprint 3)
Stephen Grace	University of East London (book sprint 2)
Alex Green	The National Archives (book sprint 2)
Edith Halvarsson	British Library (book sprint 3)
Anna Henry	Tate Gallery (book sprint 1)
Sarah Higgins	University of Aberystwyth (book sprint 3)
Jeremy Leighton John	British Library (Digital forensics)
Maggie Jones	formerly Digital Preservation Coalition (Handbook 1st edition)
William Kilbride	Digital Preservation Coalition (book sprint 1, book sprint 3, book sprint 4)
Gareth Knight	London School of Hygiene and Tropical Medicine (book sprint 2)
Sharon McMeekin	Digital Preservation Coalition (book sprint 2, book sprint 3, book sprint 4)
Paul Miller	Cloud of Data (Cloud services)
Jenny Mitcham	University of York (book sprint 3)
Laura Peaurt	University of Sheffield (book sprint 3)
Maureen Pennock	British Library (Web archiving)
Ed Pinsent	University of London Computer Centre (book sprint 1)
Virginia Power	Jisc (book sprint 1)

Vicky Stretch	Network Rail (book sprint 3)
Susan Thomas	Bodleian Library University of Oxford (book sprint 1)
Dave Thompson	Wellcome Trust (book sprint 3)
Paul Wheatley	Digital Preservation Coalition (book sprint 2, book sprint 3, book sprint 4)
Simon Wilson	Hull University Archives (book sprint 3)
Richard Wright	formerly BBC (Moving picture and sound)

Conseil d'administration

PRÉSIDENT : William Kilbride	Digital Preservation Coalition
John Chambers, Chris Fryer	Archives and Records Association
Maureen Pennock	British Library
Neil Beagrie (editor)	Charles Beagrie Ltd
Neil Grindley, Virginia Power, Matt Faber	Jisc
Emma Markiewicz, Matt Greenhall, Jane Anderson, Isobel Hunter	The National Archives

Comité consultatif

PRÉSIDENT : William Kilbride	Digital Preservation Coalition
Nancy Y McGovern	MIT Libraries
Marcel Ras	NCDD, Dutch National Coalition for Digital Preservation
Timothy Gollins	The National Archives and National Records of Scotland
Stefanie Davidson	West Yorkshire Archive Service
Sarah Higgins	Department of Information Studies, Aberystwyth University
Joy Davidson	Digital Curation Centre, University of Glasgow
Virginia Power & Matt Faber	Jisc
Anna Henry	Tate Gallery
Chris Fryer	Parliamentary Archives
Simon Tanner	King's College London
Michael Day	British Library
Susan Thomas	Bodleian Library University of Oxford
Carla Shields	Public Record Office of Northern Ireland
Marion Downie	The National Archives
Hania Smerecka	Lloyds Banking Group Archives

Révision par les pairs

Nous tenons également à remercier les 26 personnes qui ont contribué à l'examen par les pairs des projets de sections du manuel correspondant à leurs domaines de compétence et d'intérêt. Nous exprimons notre reconnaissance aux membres du conseil consultatif et également à Hugh Campbell (Public Record of Northern Ireland), Andrew Charlesworth (University of Bristol), Carey Clifford (Grosvenor Group), Lee Hibberd (National Library of Scotland), Neil Jefferies (University of

Oxford), Catherine Jones (Science and Technology Facilities Council), Naomi Korn (Naomi Korn Copyright Consultancy Ltd), Ingrid McDonald (Queensland State Archives), Laura Molloy (University of Oxford), Barbara Sierman (KB), Charlene Taylor (Worcestershire Archive & Archaeology Service), Kate Watson (Dorset County Council), Rebecca Webster (UCL Institute of Education), et Cathy Williams (The National Archives) pour leurs commentaires et suggestions qui ont contribué de manière très significative au manuel.

Références

APARSEN, 2012.*D43.1 Survey for the Assessment of Training Material/Assessment of Digital Curation Requirements*.
Disponible à l'adresse suivante: https://www.dpconline.org/docs/knowledge-base/1817-2012-02-21-aparsen-d43-1/file

DigCurV, 2013.*A Curriculum Framework for Digital Curation*.
Disponible à l'adresse suivante : https://www.digcurv.gla.ac.uk/

DPC, 2014a.*Report on the Preparatory User Consultation on the 2nd Edition of the Digital Preservation Handbook*.
Disponible à l'adresse suivante : https://www.dpconline.org/docs/miscellaneous/advice/1251-handbook-survey-response-summary-redacted/file

DPC, 2014b.*Draft Outline of the 2nd Edition of the Digital Preservation Handbook*.
Disponible à l'adresse suivante : https://www.dpconline.org/docs/miscellaneous/advice/1306-handbook-new-contents/file

Jones, M. & Beagrie, N., 2001.*Preservation Management of Digital Materials, A Handbook*. The British Library

UKWA (UK Web Archive), 2009.*8 captures (April 2008-December 2009) of the first edition of the online version of the Handbook.*
Disponible à l'adresse suivante : https://www.webarchive.org.uk/wayback/archive/20090317142605/http://www.dpconline.org/graphics/handbook/

Récapitulatif sur la préservation numérique

Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

À qui s'adresse ce chapitre ?

Administrateurs généraux (DigCurV : *Executive Lens)*, encadrants (DigCurV : *Manager Lens*) et personnels opérationnels (DigCurV : *Practitioner Lens*), organismes de financement, créateurs et éditeurs, toute personne ayant besoin d'une introduction au sujet.

Niveau de connaissance présumé

Novice.

Objet

- Fournir une vue d'ensemble stratégique et un récapitulatif à destination de la direction générale, en soulignant les grandes questions et en justifiant le financement à allouer aux tâches impliquées par la préservation des ressources numériques.
- Fournir une synthèse des réflexions actuelles sur les questions de préservation numérique.
- Faire la distinction entre les principales catégories de questions.
- Aider à clarifier l'impact des différentes questions sur les décisions à différents stades du cycle de vie des objets numériques.
- Fournir un point de départ pour de nouveaux débats et discussions au sein des organisations et avec des publics externes.

Pourquoi la préservation numérique est-elle importante ?

Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Introduction

Cette section, ainsi que la section Problèmes de préservation, est conçue comme un récapitulatif pour les novices en matière de préservation numérique. Elle est structurée en quatre sous-sections reliées entre elles. En outre, elle est étroitement liée à la section Pour commencer, qui est également conçue pour les novices en matière de préservation numérique.

La préservation numérique : le défi d'une génération

Tout objet numérique peut être concerné par la préservation numérique : nativement numérique ou numérisé, d'entreprise ou personnel, innovant ou routinier. La préservation numérique peut englober les textes et les images, les bases de données et les tableurs, les vecteurs ou les plans Rasters, les programmes et les applications, les fichiers bureautiques et les systèmes de gestion documentaire d'entreprises, le courrier électronique et les médias sociaux, les jeux, les films, la musique et le son, les domaines web et les tweets individuels. Les collections numériques peuvent provenir d'ordinateurs portables, d'ordinateurs de bureau ou de téléphones intelligents, de tablettes, de gros serveurs ou d'ordinateurs centraux. Elles peuvent être capturées au bout d'une perche à selfie ou téléportées par des capteurs à travers l'espace. Elles peuvent être créées par des caisses et des distributeurs automatiques, par des satellites et des scanners, par de minuscules puces sensibles et des réseaux massifs. Elles peuvent être stockées dans des dépôts ou des centres de données ou sur des clés USB. Il n'existe aucun objet ou système numérique qui ne soit, au moins temporairement, soumis à des questions de préservation numérique.

Invasives, changeantes et omniprésentes, les technologies numériques sont caractéristiques de notre époque. Les objets numériques sont un produit de base pour l'industrie, le commerce et les États. Ils sont fondamentaux pour la recherche, le droit et la médecine. Les industries créatives, le patrimoine culturel et les médias dépendent d'un accès fiable aux objets numériques, tandis que les familles et les amis étendent et maintiennent leurs relations grâce aux interactions numériques.

Mais les objets numériques - et les opportunités qu'ils créent - sont fragiles, même s'ils ont aussi la capacité d'être durables grâce à la reproduction. Les plateformes numériques évoluent et leurs longues chaînes d'interdépendance sont complexes et fluides. Leur longévité et leur utilité sont

menacées lorsque des contenus ou des contextes sont perdus ; l'engagement et l'exploitation sont rendus possibles lorsque les objets numériques durent. Plus les objets numériques sont importants, plus leur préservation est nécessaire : la préservation numérique protège l'investissement, capture le potentiel et transmet les opportunités aux générations futures et à la nôtre.

Nous avons déjà fait de grands progrès pour éviter une « ère d'amnésie numérique ». Il existe un nombre croissant de dépôts d'archives dans le monde entier qui peuvent se prévaloir d'une longue expérience dans la préservation des objets numériques sur plusieurs décennies (par exemple, les *UK Data Archive* fondées en 1967). Cela nous donne une large base d'expérience et des réseaux de collaboration professionnelle sur lesquels nous pouvons nous appuyer.
Il s'agit d'un défi commun à toutes les générations.

Le défi toujours émergent de la préservation numérique

La caractéristique commune des objets numériques est leur dépendance à la machine. Les informations ne sont accessibles et les fonctions ne peuvent être exécutées que par un ordinateur. À mesure que la technologie devient plus sophistiquée, cette dépendance devient une chaîne d'interdépendances de plus en plus élaborée, difficile à suivre et délicate à maintenir.

Tant que le secteur des technologies de l'information restera innovant dans la fourniture de nouveaux outils et technologies, les responsables de la préservation numérique réagiront en concevant des stratégies efficaces pour assurer la durabilité et la facilité d'utilisation des nouveaux objets numériques, de sorte que la préservation numérique restera un défi toujours renouvelé.

Pour garantir la valeur des objets numériques à long terme, nous devons en assurer l'accès, ce qui signifie que nous devons comprendre et atténuer les changements rapides dans les technologies et les organisations (voir Problèmes de préservation).

La plupart du temps, un objet numérique ne peut être bien archivé que sous forme numérique : il n'existe pas d'équivalent non numérique comme le papier qui conserverait à la fois toutes les informations essentielles et fournirait les fonctionnalités. Trop souvent, il a été nécessaire d'imprimer les objets numériques pour les archiver voire de les numériser à nouveau par la suite en raison de l'incapacité à gérer ces objets nativement numériques.

Aujourd'hui, nous disposons d'un ensemble croissant et efficace d'approches, d'expériences et de collaborations pour relever ces défis. La préservation numérique est une entreprise importante, nécessaire et réalisable, dont les premières étapes sont simples et que tout le monde peut entreprendre (voir Pour commencer).

Quel est le champ d'application ?

Ce n'est pas parce que tout pourrait faire l'objet d'une stratégie de préservation numérique que tout doit être préservé.

La question est moins de savoir ce qui peut être préservé que ce qui ne doit pas être perdu. La sélection, l'évaluation et l'élimination sont des éléments importants dans toute activité de gestion numérique. Dans le contexte d'un univers numérique en expansion, un effort déterminé pour identifier, traiter et conserver les objets numériques de valeur durable signifie d'une part que le bon objet est disponible pour les bonnes personnes au bon moment et dans le bon format, et

d'autre part que l'on identifie l'objet qui peut être activement supprimé ou négligé sans que cela ne porte à conséquence.

Les objets numériques offrent de nouvelles possibilités d'accès et d'utilisation des dépôts. Si les collections numériques existent dans un environnement qui évolue rapidement, nous devons nous attendre à ce que nos utilisateurs en fassent autant. Il est probable que les utilisateurs d'objets numériques recourent à une technologie qui n'est pas encore complètement au point, d'une manière que nous ne pouvons pas totalement anticiper, dans des endroits que nous ne visiterons peut-être jamais et à des fins que nous avons du mal à prévoir. Ainsi, toute réponse pertinente à la question « comment pouvons-nous préserver les objets numériques ? » se résumera rapidement à « que pouvons-nous faire pour garantir que ces objets numériques puissent être utilisés ? ». La planification de la préservation ne réussira que si les besoins des utilisateurs sont satisfaits.

Tout cela indique qu'il faut, dans la mesure du possible, définir la viabilité à long terme des objets numériques à un stade précoce et non tardif. Une action de préservation est nécessaire au début de la vie de l'objet, et pas toujours à sa fin. La création, la gestion et l'archivage des objets numériques ne sont plus aux extrémités opposées d'un processus mais sont intégrés tout au long de celui-ci. Par extension, la préservation n'est plus seulement une préoccupation des institutions de mémoire à long terme, mais de toute personne intéressée par l'utilisation et l'accès aux objets numériques.

Qui doit être impliqué ?

La capacité à préserver les objets numériques dépend d'un large éventail de parties prenantes. Les acteurs principaux sont les créateurs de contenu numérique, dont l'implication dans la préservation peut induire, par exemple, la prise en compte de normes en termes de format et de support, et la garantie de la disponibilité d'informations contextuelles suffisantes pour permettre leur gestion par d'autres. Les créateurs peuvent souvent ne pas être conscients de leur rôle central. Cela peut être dû à toutes sortes de raisons, mais une partie essentielle de tout effort de préservation numérique est la mise en place d'un dialogue efficace avec les créateurs d'objets numériques pour informer et défendre la valeur de leur engagement (pour eux et pour les autres).

Si les créateurs d'objets numériques ont la responsabilité de permettre l'accès à long terme, cette responsabilité est encore mieux assumée par ceux qui fournissent l'infrastructure et les environnements dans lesquels ils sont créés. Dans certains cas, il peut s'agir de la fonction d'une entreprise, avec la fourniture d'outils et de services d'entreprise adaptés à la préservation. Dans d'autres cas, la responsabilité sera assumée par des prestataires de services externes qui hébergent l'infrastructure numérique pour les clients.

La nature de la technologie numérique fait qu'il n'est pas possible de simplement transférer la gestion de la ressource à un moment donné dans le futur, sans l'avoir suffisamment gérée en amont pour en faciliter la durabilité.

Dans certains cas, les institutions géreront leur propre patrimoine numérique : les grandes institutions qui créent des objets numériques peuvent être les plus pertinentes pour les gérer à long terme, maximisant ainsi le retour sur leur investissement initial. Mais, dans d'autres contextes, des modèles coopératifs de préservation à long terme ont émergé, impliquant un certain nombre d'organisations. Des centres spécialisés et des centres d'experts ont vu le jour, offrant des solutions de préservation spécifiques pour des types d'objet numérique particuliers.

Pour certaines organisations, il peut s'avérer plus rentable de confier tout ou partie de leurs activités de préservation numérique à un tiers. Bien qu'il puisse être avantageux de sous-traiter, il est important de se rappeler que la responsabilité reste celle de l'organisation. Le personnel devra être suffisamment sensibilisé aux questions de préservation numérique, notamment en ce qui concerne les problèmes juridiques, organisationnels et contractuels, pour gérer efficacement ces contrats avec des tiers.

Toute institution qui accorde de la valeur aux ressources numériques en général doit assurer la préservation à long terme des objets numériques. Beaucoup d'institutions ont non seulement assumé ce rôle pour elles-mêmes, mais ont également pris un rôle plus large d'établissement de référence en abordant les implications pratiques de la préservation numérique.

Toutefois, la préservation numérique ne peut pas être perçue comme une préoccupation exclusive des archives, bibliothèques, musées et autres institutions patrimoniales : c'est un défi pour tous ceux qui ont intérêt à créer, utiliser, acquérir et rendre accessibles des objets numériques.

Ressources

Why Digital Preservation is Important for Everyone
https://www.youtube.com/watch?v=qEmmeFFafUs&list=PLEA69BE43AA9F7E68&index=44

Courte vidéo de la Bibliothèque du Congrès produite en 2010 pour le public non spécialisé, expliquant comment les sources d'information traditionnelles telles que les livres, les photos et les sculptures peuvent facilement survivre pendant des années, des décennies ou même des siècles, mais les éléments numériques sont fragiles et nécessitent un soin particulier pour rester utilisables. Les changements technologiques rapides affectent également la préservation numérique. Avec l'apparition de nouvelles technologies, les plus anciennes deviennent obsolètes, ce qui rend difficile l'accès aux contenus anciens (2 minutes 51 secondes).

Problèmes de préservation

Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Introduction

Cette section, ainsi que la section Pourquoi la préservation numérique est importante ?, est conçue comme un récapitulatif pour les novices en matière de préservation numérique. Elle est structurée en trois sous-sections interdépendantes, Menaces sur les objets numériques, Questions organisationnelles et Questions de ressources. Elle renvoie à des traitements plus détaillés dans d'autres sections du manuel, le cas échéant, mais elle est particulièrement proche de la section Pour commencer, qui est également conçue pour les nouveaux venus dans le domaine de la préservation numérique.

La préservation numérique peut sembler intimidante au début. Il est important de noter que les personnes qui possèdent déjà des compétences en gestion de l'information ou en technologies de l'information au sein des organisations, sont bien placées pour développer et appliquer ces compétences aux activités de préservation numérique. Toutefois, il peut être nécessaire d'apprendre au départ une nouvelle terminologie peu familière (voir Glossaire), d'élargir ses compétences et parfois de travailler différemment.

Menaces sur les objets numériques

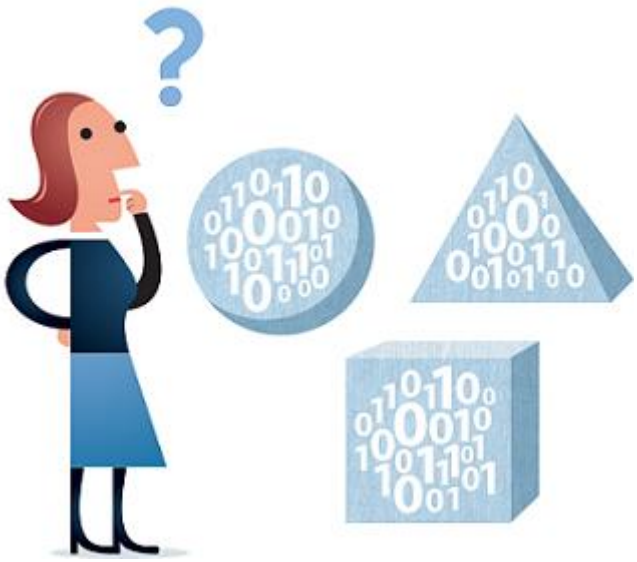


Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Conserver les données

Chaque fichier numérique est formé d'une série de zéros et de un, ou de bits (chiffres binaires). Ces flux de bits doivent être saisis et conservés dans le temps, sans perte ni dommage, pour assurer la survie des objets numériques. Toute tentative de préservation de ces bits est confrontée à une série de menaces. Les supports de stockage peuvent se détériorer au fil du temps, ce qui entraîne la corruption des fichiers. Les supports de stockage peuvent devenir obsolètes et ne plus être pris en charge par les ordinateurs modernes et les logiciels qui les interprètent et les rendent accessibles. Les bits peuvent être ignorés, abandonnés, supprimés accidentellement ou détruits par malveillance. Des supports amovibles peuvent être laissés sur une étagère et oubliés, des fichiers stockés sur un lecteur de réseau partagé peuvent être laissés sans propriétaire, ou un fournisseur de stockage *cloud* peut faire faillite.

Maintenir un processus systématique de préservation des bits reste une exigence fondamentale pour assurer la préservation numérique à long terme. Les supports de stockage doivent être contrôlés et actualisés (voir Stockage). La redondance doit être introduite en répliquant ou en sauvegardant les fichiers, via une diversité dans les technologies dépendantes et pour éviter un désastre qui interviendrait sur un seul lieu géographique (voir Stockage). Des empreintes numériques doivent être générées et fréquemment recalculées pour identifier toute perte et garantir que l'intégrité des bits peut être vérifiée de manière efficace et automatisée (voir Intégrité et empreintes). Les emplacements dans lesquels les objets numériques sont stockés doivent être soigneusement enregistrés, et la responsabilité de leur préservation doit être définie.

Conserver le sens des données

La reconstruction des informations codées dans un flux de bits nécessite par exemple un logiciel informatique conçu pour restituer, manipuler, analyser ou plus généralement agir sur l'encodage ou le format particulier des données. Au fil du temps, les encodages (ou formats de fichiers) peuvent changer et les applications logicielles qui interagissent avec eux peuvent devenir obsolètes. Bien que ce soit peu fréquent pour des formats de fichiers bien connus, les formats de

fichiers moins bien utilisés peuvent devenir obsolètes avec le temps, car le logiciel qui les restitue n'est plus maintenu (voir Formats de fichiers et normes).

La compréhension de la technologie dont dépendent certains objets numériques permet de prendre les mesures appropriées pour assurer leur préservation. Un processus de planification de la préservation bien réfléchi peut aboutir à la migration des fichiers numériques d'un format à un autre, à l'émulation de logiciels obsolètes ou à l'utilisation d'autres applications logicielles pour la restitution des données (voir Mesures de préservation). Chacune des options présente ses propres avantages et inconvénients, qui doivent être évalués avec soin, éventuellement au cas par cas (voir Planification de la préservation).

Si l'obsolescence des formats de fichiers n'a pas toujours été perçue comme un risque majeur, des subtilités délicates demeurent pourtant. Il est peut-être possible de trouver une méthode pour restituer un ancien format de fichier (peut-être en émulant un logiciel obsolète), mais quelle est la fiabilité de cette restitution ? Est-il légal d'exécuter le logiciel et combien cet effort complexe coûtera-t-il au conservateur et à l'utilisateur ?

Maintenir la confiance dans les données

Les documents numériques ont la caractéristique d'être instables dans le temps, d'être modifiables ou modifiés facilement, d'être endommagés par une défaillance du support ou décodés de manière peu fiable ou inexacte par un logiciel de restitution en informations lisibles par l'homme. Pour qu'un utilisateur final ait confiance dans le résultat d'une opération de préservation numérique, il faut qu'il prenne soigneusement en considération l'ensemble du cycle de vie des objets numériques et qu'il sache quelle personne ou quel outil a interagi avec eux au fil du temps. Les systèmes de gestion de l'information doivent pouvoir établir des liens avec les informations contextuelles essentielles concernant les processus métier de l'entité créatrice. L'authenticité et l'intégrité des ressources numériques peuvent être tout aussi importantes dans d'autres secteurs. Par exemple, les universitaires devront avoir l'assurance que les références qu'ils citent resteront les mêmes au fil du temps ; les tribunaux devront être assurés que l'objet numérique peut résister aux exigences de preuve juridique ; les ministères pourraient avoir des exigences juridiquement contraignantes en matière d'authenticité, etc. Ce problème recoupe à la fois des questions juridiques et organisationnelles, et il se peut qu'il soit mieux résolu au cas par cas plutôt que par des procédures génériques.

L'application de techniques d'intégrité des données et le maintien de journaux d'audit peuvent donner l'assurance qu'un objet numérique est resté intègre (sauf par une action de préservation nécessaire) depuis son dépôt dans un système d'archivage (voir Intégrité et empreintes, et Sécurité de l'information). En somme, son authenticité pour un utilisateur peut dépendre essentiellement de la fiabilité de l'organisme de préservation. Il sera crucial de maintenir des processus de préservation de haute qualité basés sur les meilleures pratiques actuelles et validés par un audit et une certification appropriés (voir Audit et certification).

Conserver le contexte des données et ses dépendances

La signification des informations numériques peut dépendre d'informations supplémentaires qui peuvent avoir été implicites dans le contexte où elles ont été créées ou utilisées à l'origine, et sont moins claires lorsqu'elles sont réexaminées ultérieurement. L'identification, la compréhension et la saisie d'informations contextuelles pertinentes peuvent être essentielles à la réussite des efforts de préservation. Cela peut être aussi simple que de saisir les unités de mesure utilisées dans un tableur, l'échelle d'une carte ou le point d'origine d'un dessin CAO. Comme l'information

numérique est sans cesse créée de manière plus complexe et interconnectée, il peut être nécessaire de conserver la place d'objets numériques particuliers dans un contexte plus large de ressources d'information associées. Le contenu qui semble simple et autonome peut en fait dépendre de fichiers connexes, de polices de caractères référencées et peut comporter des pointeurs vers des informations connexes sur le web. Une simple page web peut avoir été générée dynamiquement à partir de données provenant de différents endroits sur l'Internet.

La compréhension des données, de leur utilisation, de leurs dépendances et de leur contexte permettra de les capturer pour les conserver de manière appropriée et de les documenter de manière suffisamment explicite pour que le contenu intellectuel puisse être conservé et compris à l'avenir (voir Métadonnées et documentation).

Agir en temps opportun

Il peut être crucial de donner la priorité aux activités de préservation numérique et de les appliquer à temps, non seulement pour éviter les pertes, mais aussi pour garantir une utilisation optimale de ressources limitées. Lorsque la possibilité d'intervenir tôt dans le cycle de vie existe, les objets numériques peuvent être préparés pour mieux survivre à l'avenir. Le choix du format de fichier, la saisie de la documentation critique ou la description des relations clés dans les métadonnées peuvent nécessiter un petit investissement au départ, mais peuvent permettre de réaliser des économies considérables plus tard (voir Création d'objets numériques). Lorsque cela n'est pas possible et que les risques pour les données ont été identifiés, définir le meilleur moment pour appliquer des mesures de préservation n'est pas forcément évident. Une intervention précoce pour éviter l'obsolescence technologique peut donner une plus grande confiance dans la durabilité à long terme, mais avec le risque qu'en fin de compte l'intervention ne soit pas nécessaire et que les ressources soient gaspillées. Une action juste à temps peut réduire au minimum les activités inutiles, mais accentue l'effort requis pour faire des recherches sur une technologie obsolète nécessitant des connaissances spécifiques qui ne sont plus d'actualité. Les mesures appropriées doivent donc être prises au cas par cas.

Faire face au déluge de données

Selon les recherches effectuées par David Rosenthal, le taux de création de données augmente d'environ 60 % par an, les évolutions en matière de stockage de données permettent une expansion d'environ 25 % par an et les budgets des centres de données augmentent d'environ 2 % par an (Rosenthal, 2014). Si cette situation exerce des pressions stimulantes sur les politiques de sélection et autres décisions organisationnelles, elle pose également des questions d'ordre technologique. Les processus de préservation simples qui fonctionnent efficacement à un niveau ne s'adapteront pas aisément à de très grands volumes de données ou des fichiers individuels très volumineux. La technologie et la compréhension du travail à grande échelle progressent rapidement, avec une expertise croissante pour le traitement de grandes collections audiovisuelles, de données de recherche et d'archives sur le web (voir Préservation du contenu spécifique). Certains dépôts d'archives sont encore confrontés à des défis importants pour développer et maintenir des architectures et des procédures évolutives traitant des quantités croissantes de données. Il ne faut pas sous-estimer les défis techniques et managériaux liés à l'acquisition, à la gestion et à la fourniture d'un accès aux objets numériques à cette échelle. Il peut être important de se rappeler que la sélection, l'évaluation et l'élimination sont des composantes importantes de toute activité de gestion numérique.

Questions organisationnelles

Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Si les questions technologiques peuvent être complexes, il existe également de nombreux défis liés aux questions organisationnelles. Il s'agit notamment de la manière dont la préservation numérique est organisée et assurée, ou de la manière dont ces responsabilités évoluent dans le temps et au cours du cycle de vie des objets numériques. Les défis de la préservation numérique sont communs à toutes les organisations, mais chaque contexte organisationnel est différent. Il est essentiel de déterminer les facteurs organisationnels et d'adapter des solutions pratiques pour répondre à ces besoins. Il n'existe pas d'approche unique de la préservation numérique.

La création, la préservation et l'accès aux objets numériques sont largement distribués. Par conséquent, il est de plus en plus nécessaire de dépasser les limites des organisations individuelles, voire des pays, afin de maximiser les avantages de la technologie, d'aborder les problèmes communs et de relever les défis au meilleur coût.

En interne ou en externe ?

La décision de confier tout ou partie de la préservation numérique à un tiers ou à un service interne, ou peut-être à une combinaison des deux, est souvent complexe. La préservation numérique peut être entreprise en interne si le personnel et l'infrastructure sont suffisants, mais l'externalisation de certaines activités ou de certains services de soutien peut être rentable et permettre de tirer parti des capacités et des moyens internes.

L'externalisation de tâches ou de services spécifiques à partir d'un entrepôt numérique n'est en aucun cas un phénomène nouveau. Les entrepôts numériques sous-traitent certaines de leurs opérations depuis des décennies. Il est essentiel de disposer et de conserver des connaissances suffisantes pour être en mesure de préparer des spécifications efficaces et de contrôler les performances. Le travail sous-traité doit pouvoir être facilement vérifié et sa qualité contrôlée, ce qui est rendu possible par une conception soigneuse des spécifications et par la fourniture de rapports par le tiers. Le coût sera clairement un élément clé pour décider de sous-traiter ou non la préservation numérique, mais d'autres facteurs doivent être pris en compte, comme les questions

juridiques. Par exemple, les dispositions légales relatives à la vie privée ou à la confidentialité peuvent influencer l'opportunité de sous-traiter ou non. Les avantages et les inconvénients de chaque option devront être équilibrés à la lumière de la mission et des responsabilités de chaque organisation (voir les rubriques Achats et recours à des tiers et Services *cloud*).

Collaboration

Les problèmes de préservation numérique auxquels sont confrontés toutes les organisations et tous les secteurs se recoupent largement. Il est donc logique de mettre en commun les compétences et l'expérience. Il existe des raisons impérieuses et, dans certains cas, des injonctions politiques, pour s'engager dans une plus grande collaboration au sein des organisations et entre celles-ci afin d'affronter et de surmonter efficacement les défis de la préservation numérique.

La plupart des organisations reconnaissent volontiers les avantages d'une collaboration accrue, mais indiquent également les difficultés potentielles qui peuvent se présenter sous la forme de différences de programmes, de calendriers ou de mécanismes de financement. Néanmoins, il est souvent possible de collaborer dans des domaines spécifiques ou avec des niveaux d'intensité différents qui atténuent ces difficultés potentielles. Certaines des initiatives les plus en vue et les plus réussies de ces derniers temps dans le domaine de la préservation numérique ont été de nature collaborative (voir Collaboration).

Changement organisationnel

Le monde numérique moderne est un lieu de changements technologiques et organisationnels rapides. Les organisations se réorganisent en interne, fusionnent ou cessent leurs activités avec une fréquence croissante. La préservation numérique est une activité de long terme et la probabilité qu'elle soit affectée par des changements organisationnels augmente avec le temps. Un dépôt peut être affecté non seulement par des changements au niveau de son organisation mère, mais aussi par des changements au niveau de ses principaux déposants et utilisateurs, fournisseurs ou collaborateurs. Le changement organisationnel est donc un risque majeur à gérer (voir Gestion des risques et du changement).

Structures organisationnelles

La nature de la technologie et les dépendances dans la préservation des objets numériques sont telles qu'elles ont des implications sur les structures organisationnelles. De nombreuses activités sont convergentes, par exemple les décisions concernant l'acquisition et la préservation doivent être prises judicieusement en même temps. Les structures organisationnelles devront franchir certaines frontières afin de tirer parti de toute la gamme des compétences et de l'expertise requises pour les objets numériques. L'attribution de la responsabilité de la préservation des objets numériques acquis et/ou créés au sein d'une organisation nécessitera inévitablement la participation conjointe du personnel de différentes parties de l'organisation. Cela peut présenter des difficultés, à moins d'être soutenu par une vision d'entreprise forte qui peut être communiquée au personnel (voir Collaboration, Sensibilisation et Formation et développement du personnel).

Rôles et responsabilités

Il existe certains dépôts d'archives qui assument la responsabilité de domaines ou de formats spécifiques. Au Royaume-Uni, par exemple, le *UK Data Service* a la responsabilité de certaines données de recherche en sciences sociales, tandis que la *National Sound Archive* de la *British Library* est en charge de sa collection d'enregistrements sonores. Chaque dépôt devra tenir

compte de sa propre politique de collecte et du paysage plus large des institutions de collecte et des attributions dans lesquelles il s'inscrit.

L'environnement numérique exige l'engagement d'un large groupe de parties prenantes. L'approche de la préservation numérique basée sur le cycle de vie, qui est préconisée dans ce manuel, a des implications importantes sur la manière dont les organisations responsables de la préservation à long terme doivent interagir et collaborer avec les créateurs, les éditeurs et autres intermédiaires, et entre elles.

Les créateurs d'objets numériques doivent être capables de comprendre les implications de leurs actions en termes de pérennité à moyen et long terme pour l'objet numérique qu'ils créent. Qu'il s'agisse d'un document créé dans le cadre des activités quotidiennes du service, d'une copie numérique d'un objet analogique ou d'une ressource « nativement numérique », des conseils et un soutien, ainsi qu'une infrastructure technique et organisationnelle appropriée, contribueront à améliorer considérablement les perspectives de gestion et de préservation (voir Création d'objets numériques).

Sélection

L'énorme quantité d'informations produites sous forme numérique, leur qualité variable et les contraintes de ressources auxquelles sont soumis ceux qui prennent la responsabilité de préserver l'accès à long terme, rendent la sélection inévitable si l'objectif est de préserver un accès continu.

Dans l'environnement numérique, la non-sélection pour la préservation est quasi systématiquement synonyme de perte, même si cet aspect est ensuite considéré comme valable.

Dans les cas où il peut y avoir plusieurs versions, il faut décider quelle version est la meilleure pour la préservation, ou s'il faut en choisir plusieurs. L'échantillonnage des ressources dynamiques, par opposition à la tentative de sauvegarde de chaque changement, peut-être la seule option envisageable mais peut avoir de graves répercussions si l'échantillonnage n'est pas entrepris dans un cadre bien défini et en tenant dûment compte des besoins connus, contemporains et futurs, des utilisateurs.

Il convient également de tenir compte, lors de la sélection, du niveau de redondance nécessaire pour assurer la préservation numérique. Il faut bien comprendre qui assumera cette responsabilité et pour quelle période de temps. Dans le cas contraire, même si plusieurs copies sont stockées dans divers entrepôts, celles-ci pourraient, pour de multiples raisons, cesser d'assurer la maintenance de l'objet numérique à un moment donné (voir aussi Acquisition et évaluation).

Concilier sécurité et accès

Il y a toujours eu un lien étroit entre la préservation et l'accès. Les dépositaires doivent s'assurer que leurs objets numériques sont sûrs et sécurisés, mais la plupart d'entre eux fournissent également un accès à une large gamme d'utilisateurs. L'accès par des utilisateurs réels peut constituer un fil directeur précieux pour la conception des services de préservation, en permettant d'éviter des actions inutiles, mais aussi en validant et en introduisant des retours d'expérience réguliers.

De nombreux types d'objets numériques sélectionnés pour la préservation à long terme peuvent contenir des informations confidentielles et sensibles qui doivent être protégées pour garantir

qu'elles ne sont pas accessibles à des utilisateurs non autorisés. Dans d'autres cas, l'entrepôt peut être soumis à des obligations légales ou réglementaires qui en affectent l'accès. Il peut y avoir des tensions entre ces deux rôles et il peut donc être nécessaire de trouver un équilibre entre la sécurité et la facilité d'accès (voir Accès et Sécurité de l'information).

Respect de la législation

Les questions juridiques ne sont pas simples en matière de préservation numérique. Il existe souvent des copies multiples et des versions dérivées d'objets numériques, et il peut y avoir des logiciels et des métadonnées qui leur sont associés, provenant de différentes sources. Le contenu numérique est généré par un groupe plus large de créateurs et intègre des formats et des droits de propriété intellectuelle (DPI) plus divers que ceux qui s'appliquent dans le monde analogique. En outre, le droit est souvent à la traîne par rapport à l'évolution technologique et aux besoins de préservation numérique. Voici quelques-unes des principales questions juridiques qui touchent les dépositaires dans la collecte, la préservation et la fourniture d'accès aux objets numériques :

- toute exigence juridique en termes de gestion, de préservation et d'accès imposée au dépôt et à son organisation mère, par les donateurs et les financeurs par le biais de contrats et d'accords ou par la législation du gouvernement (par exemple, l'accessibilité, la disponibilité, la sécurité des informations, la préservation, l'audit et la conformité, les documents publics, le dépôt légal, etc.) ;
- les obligations légales relatives aux droits des tiers sur les documents numériques détenus par le dépôt (par exemple, droits d'auteur, protection des données) ;
- les éléments juridiques relatifs à toute relation entre un référentiel et un ou plusieurs fournisseurs tiers (par exemple, les conditions des contrats de service et les accords de niveau de service).

Pour obtenir des conseils et des ressources supplémentaires permettant de traiter ces questions et de gérer les risques associés, consultez les sections Respect de la réglementation et Achats et recours à des tiers.

Questions de ressources

Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Budget et coûts

Il est complexe d'isoler le coût de la préservation numérique des autres dépenses d'organisation, et il ne devrait pas l'être. La préservation numérique consiste essentiellement à préserver l'accès au fil du temps et, par conséquent, les coûts de tous les éléments du cycle de vie numérique sont à prendre en compte. Dans ce contexte, même les coûts de création des objets numériques sont à intégrer dans la mesure où ils peuvent inclure des éléments de coût qui faciliteront leur préservation à long terme (voir Création d'objets numériques).

La capacité à employer et à former du personnel possédant les compétences appropriées est rendue plus difficile par la rapidité de l'évolution technologique et l'éventail des compétences nécessaires. Elle est également limitée par les contraintes de ressources des organisations qui peuvent avoir besoin de gérer des collections traditionnelles et numériques croissantes sans ressources supplémentaires.

Néanmoins, l'exercice de calcul des coûts, aussi complexe soit-il, est un exercice utile et nécessaire pour établir des pratiques rentables et un modèle économique fiable. Le coût du travail nécessaire à la préservation numérique sera de loin le plus important et comprend non seulement des experts spécialisés, mais aussi des proportions variables d'efforts de la part de nombreux membres du personnel tels que l'administration, la gestion, le soutien informatique, les conseillers juridiques, etc.

D'autres questions importantes ayant une incidence sur les coûts sont la mission et les objectifs de l'organisation, notamment le type et la taille des collections, le niveau de préservation engagé, la quantité et le niveau d'accès requis, et le calendrier d'action proposé. Ces questions sont examinées en détail dans la section Analyse de rentabilité, avantages, coûts et impact.

La relation entre les coûts et les stratégies et activités institutionnelles telles que la collaboration, l'achat et le recours à des tiers, le respect de la législation, la formation et le développement du personnel ou les normes et les bonnes pratiques est également abordée dans les sections correspondantes du manuel.

Personnel et compétences

La préservation numérique implique toute une série de compétences et de rôles organisationnels. En règle générale, la préservation numérique fait appel à une série de compétences qui ne sont pas combinées habituellement. Cela signifie que les grandes organisations devront probablement constituer des équipes pluridisciplinaires, tandis que dans les petites organisations, il sera nécessaire de s'appuyer sur des équipes ou des ressources partagées.

Il y a trois questions principales à considérer en ce qui concerne les effectifs et les compétences :

- Premièrement, bien qu'il y ait eu des améliorations considérables ces dernières années, l'enseignement de la préservation numérique est souvent en retard par rapport aux meilleures pratiques actuelles ou est entièrement théorisé dans le cadre de programmes de gestion de l'information qui s'avèrent plus pertinents pour les nouveaux arrivants dans la profession. Les personnes ayant des compétences et une expérience pratique sont donc très demandées et le personnel peut être difficile à recruter.

- Deuxièmement, les descriptions de poste peuvent être difficiles à rédiger, surtout lorsque les agences partent effectivement de zéro pour définir un nouveau rôle. Dans cette optique, un certain nombre de projets de recherche a tenté de décrire les compétences génériques nécessaires à la préservation numérique, en partant de l'hypothèse que différentes compétences sont requises à différents niveaux d'une organisation. Des outils tels que le cadre de compétences DigCurV allié à la section *Vacancies* de *Digital Preservation Coalition* peuvent être très utiles pour décrire les nouveaux rôles. Les grandes organisations dotées d'équipes multidisciplinaires peuvent être en mesure de recruter sur des postes qui sont des variantes « numériques » de catégories professionnelles existantes telles qu'archiviste, bibliothécaire ou gestionnaire de documents, mais pour la plupart des organisations, de nouveaux types de postes doivent être créés.

- Enfin, le personnel travaillant dans le domaine de la préservation numérique signale fréquemment la nécessité de rester à la page pour développer sa carrière. Étant donné que la technologie et les besoins des utilisateurs évoluent avec le temps, le personnel chargé de répondre à ces nouvelles exigences devra trouver des moyens de mettre à jour ses compétences en permanence, par exemple en organisant des réunions d'information pour les spécialistes et en créant des réseaux professionnels (voir Formation et développement des compétences du personnel).

Équipements

Une préservation numérique efficace nécessite certaines installations ou infrastructures de base, généralement de nature technologique, sur lesquelles peuvent reposer les processus de travail opérationnels et le traitement des objets numériques. Même si ces installations peuvent être rudimentaire, ou du moins de petite taille, lorsqu'une organisation fait ses premiers pas dans la préservation numérique, l'accélération des opérations pour traiter de grandes quantités de données nécessitera des investissements considérables dans les installations nécessaires pour les mettre en œuvre.

Stockage

Avec l'exigence caractéristique de réplication des données préservées pour éviter leur perte, le matériel de stockage reste l'une des installations de préservation numérique les plus importantes. La technologie de stockage a évolué rapidement au cours des dernières décennies. Les archives ont

largement utilisé des supports tels que les CD ou les DVD pour le stockage à long terme, mais les développements rapides des supports magnétiques ont apporté un stockage rapide et fiable qui a rendu les supports portables superflus. Les systèmes de stockage des entreprises fournissent désormais de grands volumes de stockage à un coût raisonnable. Bien que leur durée de vie soit limitée, généralement de 4 à 8 ans environ, ils sont faciles à surveiller et à remplacer lorsqu'ils arrivent en fin de vie (voir Stockage).

Les organisations peuvent également envisager des services *cloud* pour « louer » l'infrastructure de préservation. La flexibilité du *cloud* permet des tests et un pilotage relativement rapides et peu coûteux. Les services *cloud* peuvent fournir une réplication facile et automatisée vers plusieurs endroits et un accès à un stockage numérique géré par des professionnels, ainsi qu'à un contrôle d'intégrité. Les entrepôts numériques peuvent ajouter l'accès à des outils, des procédures, des processus métier et des accords de service dédiés, fournissant un système d'entrepôt numérique adapté aux exigences de la préservation numérique par l'intermédiaire de fournisseurs spécialisés (voir Services *cloud*).

Systèmes d'entrepôt numérique

La plupart des exigences de base pour la préservation des objets numériques est fournie de manière automatisée par des systèmes de préservation numérique dédiés, ou des entrepôts numériques fiables. Une application de l'entrepôt identifiera de manière unique chaque objet numérique qui y est placé. Elle gérera le stockage de cet objet, identifiera ses caractéristiques et aidera un gestionnaire de l'entrepôt à planifier sa préservation. Elle facilitera également l'accès à l'objet. Si la préservation de base peut être assurée de manière ponctuelle à petite échelle, une application spécialisée est essentielle pour gérer efficacement les objets numériques dans le temps. OAIS est un modèle de haut niveau pour les fonctions requises par un entrepôt d'archives (voir Audit et certification pour plus d'informations sur la certification des dépôts numériques fiables et Outils pour les systèmes et composants de dépôt d'archives).

Calcul haute performance

L'augmentation des volumes de données nécessite non seulement plus de stockage mais aussi une plus grande puissance de calcul. La caractérisation et l'évaluation des caractéristiques techniques des données, l'indexation des données pour permettre la recherche et l'accès, la vérification de l'intégrité et une variété d'autres tâches exigent des performances de calcul considérables. Ceux qui traitent ces masses de données, qu'il s'agisse de données de recherche ou d'archives web, se tournent généralement vers le calcul haute performance, et vers des technologies telles qu'Apache Hadoop fonctionnant sur des clusters de matériel pour répondre à ce besoin.

Laboratoire de préservation numérique

Un certain nombre de grandes organisations ont développé des laboratoires dans lesquels un ensemble de technologies anciennes et nouvelles peuvent être appliquées pour la stabilisation ou l'extraction de données à partir de supports obsolètes. Ces laboratoires ont été promus par des organisations travaillant avec des collections numériques particulières. Des lecteurs spécialisés pour la lecture de supports magnétiques, des robots pour le traitement d'un grand nombre de disques optiques et des bloqueurs d'écriture pour permettre l'accès aux disques durs sans changer les bits en cours de route, sont quelques-uns des équipements qui peuvent être utilisés. Les entreprises de récupération de supports proposent une autre approche qui semble préférable

dans le cas de gros volumes, mais avec un contrôle moindre du processus et la nécessité de déplacer les supports hors site (voir Criminalistique numérique).

Ressources

How Toy Story 2 Almost Got Deleted: Stories From Pixar Animation: ENTV
https://www.youtube.com/watch?v=8dhp_20j0Ys
Histoire divertissante et informative sur la façon dont *Toy Story 2* a presque été supprimé des ordinateurs de *Pixar Animation* pendant le tournage du film et comment le film a été sauvegardé sur l'ordinateur familial d'une mère (2 minutes 26 secondes)

Références

Rosenthal, D., 2014. *Talk "Costs: Why Do We Care?", DSHR's Blog*, Tuesday November 18 2014. Disponible à l'adresse suivante : http://blog.dshr.org/2014/11/talk-costs-why-do-we-care.html

Pour commencer

Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Introduction

Cette section s'adresse à vous si vous n'avez pas encore commencé le processus de préservation numérique ou si vous venez de le faire. Elle fournit une introduction rapide à un certain nombre d'approches qui vous aideront à démarrer, à utiliser d'autres sections du manuel et à renforcer votre confiance et vos compétences.

Cette section est issue des ateliers « *Getting Started in Digital Preservation* » organisés par *Digital Preservation Coalition*. Elle soutient l'« apprentissage par la pratique » et suppose un niveau minimal de connaissances préalables. Au fur et à mesure de votre progression, vous pourrez puiser dans les ressources et les études de cas, dans d'autres sujets et sections du manuel, et dans la section Glossaire pour les termes peu familiers.

La préservation numérique peut décourager au premier abord. Toutefois, elle peut vous aider à déterminer les compétences et les ressources dont vous disposez, ainsi que les objets que vous souhaitez préserver. Ainsi, il faut commencer avec ce que vous savez plutôt qu'avec ce que vous ne savez pas. La première étape de la préservation numérique consiste presque toujours à entreprendre une évaluation rapide. Celle-ci comportera deux ou trois volets :

- connaître la capacité pratique de votre organisation ;
- comprendre les objectifs et les missions de l'organisation ;
- acquérir une base de connaissances sur les objets numériques en question.

Apprendre à connaître votre organisation et vos données

Création d'un registre d'objets numériques

Dans le cadre d'une évaluation rapide, il est essentiel de comprendre la nature et l'étendue de vos collections numériques. Un registre d'objets numériques peut être d'une très grande utilité pour évaluer l'étendue et l'importance de votre collection, identifier les priorités et planifier les actions de préservation numérique. Une évaluation générale de la collection aidera à dresser une carte plus détaillée par la suite : un audit complet et détaillé pourrait prendre beaucoup de temps. Il est donc conseillé, dès les premières étapes, de privilégier la simplicité au niveau du registre d'objets numériques.

Posez-vous les questions suivantes :

- Quel est le sujet de la collection ?
- D'où vient-elle et quelle est sa fonction ?
- Où est-elle stockée et quels types de supports sont utilisés ?
- Pourquoi est-elle conservée ?
- Qui en est responsable ? Qui sont les utilisateurs ? De quels sujets traitent les données ?
- Comment accède-t-on aux données ?
- Comment les données sont-elles susceptibles de changer et de se développer dans un avenir proche ?

Évaluation de l'état de préparation de votre organisation

La maturité organisationnelle est un autre facteur à prendre en compte. La *National Digital Stewardship Alliance* (NDSA, 2013) aux États-Unis a recommandé l'utilisation d'un modèle simple à quatre niveaux pour aider les organisations à comprendre et à améliorer leurs capacités techniques en matière de préservation numérique. Les quatre niveaux sont les suivants :

- Niveau 1 : protéger vos données
- Niveau 2 : connaître vos données
- Niveau 3 : surveiller vos données
- Niveau 4 : réparer vos données

Ces « niveaux de préservation » se veulent progressifs et servent à mesurer la maturité en fonction de cinq éléments : le stockage, l'intégrité des fichiers, la sécurité des informations, les métadonnées et les formats de fichiers. La capacité d'une organisation à entreprendre la préservation numérique est indiquée par son niveau de maturité pour ces cinq composantes. Des modèles de maturité plus complets sont disponibles, tels que le modèle de maturité de la capacité de préservation numérique (Dollar et Ashley, 2014), dans le cas où une analyse plus approfondie de la maturité organisationnelle est nécessaire.

Premières étapes pour sécuriser vos données

Cette section fournit un aperçu des premières mesures à prendre pour sécuriser vos données une fois que vous avez évalué l'état de préparation de votre organisation et compilé des informations de base sur vos données. Les étapes suivantes sont essentielles pour garantir un niveau minimum de préservation lorsque vous recevez une nouvelle collection d'objets numériques. Cette étape est généralement appelée « préservation des bits » . Il s'agit littéralement de préserver les flux de chiffres binaires, ou bits, qui composent vos fichiers numériques (sans préserver les moyens de décoder les bits en informations significatives).

Vérification rapide à la réception

Lorsqu'une nouvelle collection d'objets numériques est obtenue de la part d'un fournisseur, il est essentiel de s'assurer que ce qui a été reçu correspond à ce qui est attendu. Selon la source de l'objet, il peut être possible de demander de nouvelles copies des fichiers manquants ou de mauvaise qualité. Une fois ces vérifications effectuées et les demandes de remplacement correspondantes demandées, le processus a de meilleures chances d'aboutir.

Les principales tâches sont les suivantes :

- Rechercher les virus et les logiciels malveillants pour s'assurer qu'aucun élément indésirable ne se trouve dans la collection. Garder éventuellement la collection « en quarantaine » jusqu'à sa vérification.

- Vérifier que tous les fichiers attendus sont présents. Si l'objet est accompagné d'un manifeste, vérifier les fichiers par rapport à celui-ci.
- Ouvrir une sélection aléatoire de fichiers pour vérifier leur intégrité et/ou les niveaux de qualité attendus.
- Demander rapidement le remplacement de tout fichier endommagé ou manquant, si possible.

Création d'une liste de fichiers vérifiables

Afin de vérifier au fil du temps que vos fichiers numériques sont préservés, il est d'abord nécessaire d'enregistrer exactement quels sont les fichiers en votre possession. Il est donc important de créer une liste vérifiable de fichiers de chaque collection. Ces listes devraient probablement contenir des informations telles que les noms de fichiers, les emplacements et les tailles, les types de formats et les empreintes. Une empreinte numérique est une courte chaîne alphanumérique qui représente le contenu d'un fichier et qui agit comme une « empreinte digitale humaine » permettant une comparaison dans le temps. Une fois la liste créée, un processus simple permet de vérifier que tous les fichiers sont présents et intacts, à tout moment dans le futur (voir la vérification de l'intégrité ci-dessous).

Différents outils peuvent être utilisés pour générer automatiquement ces données ; ils sont communément appelés outils de caractérisation. Par exemple, vous pouvez utiliser PRONOM des Archives nationales du Royaume-Uni (un registre des formats de fichiers et de leurs comportements) et DROID (un outil qui utilise PRONOM pour analyser les fichiers sur un système). En disposant d'une liste des formats, versions et quantités de fichiers de vos collections, vous pourrez démontrer auprès de la direction générale le soutien et les ressources qui seront nécessaires pour effectuer un travail efficace et pérenne. Ces informations permettront également de mettre à jour et d'enrichir votre registre de ressources numériques. L'éventail des formats utilisés doit être consolidé afin de réduire au minimum la duplication et d'éliminer les formats problématiques. Ce processus est connu sous le nom de « standardisation ».

Les tâches principales sont les suivantes :

- Générer une liste de fichiers vérifiables ;
- Mettre à jour le registre des objets numériques.

Sécurisation de vos fichiers : création de copies

Quelle que soit la qualité de votre stockage numérique, votre objet numérique sera toujours exposé à des risques de détérioration, de dégradation ou de suppression accidentelle. Faire plus d'une copie de vos objets numériques et utiliser plus d'un type de solution de stockage permet d'atténuer divers risques de préservation numérique.

Les tâches principales sont les suivantes :

- Conserver (au moins) une copie facilement accessible sur un disque non amovible. Vous devrez réexaminer régulièrement votre objet pour vous assurer de son intégrité, donc le garder accessible facilitera cette tâche.
- Faire (au moins) une copie supplémentaire, si nécessaire sur un support de stockage moins accessible, mais moins coûteux, tel que la bande magnétique.
- Conserver une copie dans un lieu géographique différent des autres afin de minimiser les risques.

Révision et inspection : vérification de l'intégrité

En réexaminant vos documents numériques régulièrement (par exemple tous les 6 mois), vous pouvez vous assurer qu'aucun dommage ou perte accidentelle ne s'est produit. Si c'est le cas, vous pouvez récupérer les fichiers problématiques à partir des copies ou des sauvegardes que vous avez faites précédemment. Les futurs contrôles d'intégrité généreront de nouvelles empreintes numériques pour les fichiers de vos collections. Si elles ne correspondent pas à celles créées à l'origine, il y a eu perte ou dommage de bits.

Les principales tâches sont les suivantes :

- Réexaminer fréquemment votre collection, recalculer les empreintes, identifier les fichiers endommagés.
- Récupérer les copies des fichiers endommagés et les réparer si nécessaire.
- Effectuer des tests de récupération des données sauvegardées par des services tiers, afin de s'assurer que les sauvegardes sont effectuées comme convenu.

Détails de vos processus

Dès le début de la création d'une collection numérique, il est important de documenter autant que possible les éléments d'une collection, les outils et les processus de travail. Cette documentation est une composante importante des métadonnées techniques et descriptives. Ces informations doivent être conservées pour des raisons de pérennité. Comme pour tout projet, la continuité du personnel peut poser problème. Lorsque le personnel change, il emporte souvent des connaissances et des compétences essentielles.

Prochaines étapes

Après avoir fait vos premiers pas dans la préservation numérique, quelles sont les prochaines étapes ? Cela dépendra évidemment de vos propres exigences et priorités. Ce tableau fournit un certain nombre de suggestions ; d'autres sections du manuel vous aideront à les mettre en œuvre :

Étapes suivantes
Développez la sensibilisation et la communication, la compréhension des risques, l'analyse de rentabilité, avantages, coûts et impacts.
Établissez une stratégie et des politiques de préservation pour l'organisation. En plus de garantir une approche cohérente de la préservation, il peut être utile d'obtenir l'adhésion de l'ensemble de l'organisation, et en particulier de la direction générale.
Mettez en place un entrepôt numérique. Des solutions et des outils techniques, soit sur l'infrastructure informatique locale, soit proposés sous forme de services *cloud*, vous aideront à comprendre, gérer et préserver vos objets numériques à long terme.
Établissez votre offre de stockage durable et vos plans de préservation et d'action à long terme
Réexaminez et élargissez vos audits de collections : - Caractérisez plus en détail les collections prioritaires ; - Mettez à jour périodiquement les audits, le cas échéant.
Créez un groupe de travail sur la préservation numérique. Une préservation numérique efficace nécessite souvent l'adhésion de nombreux services au sein d'une organisation. Un groupe de travail représentatif peut s'avérer essentiel pour faire avancer les choses de manière coordonnée.
Mettez en place la formation et le développement des compétences du personnel'.
Établissez un réseau professionnel et des collaborations. Adhérez à une organisation de

préservation numérique telle que la *Digital Preservation Coalition*.
Tenez-vous au courant de l'actualité: - Les listes de diffusion pour la préservation numérique comprennent la liste d'annonces de préservation numérique sur JiscMail, et la liste *digipres* axée sur les États-Unis[4] ; - Un blog hebdomadaire de *DP News* sélectionne des tweets et des liens d'actualité récents sur la préservation numérique[5] ; - Parmi les revues axées sur la préservation numérique, on peut citer l'*International Journal of Digital Curation* et D-Lib ; - Parmi les événements axés sur la préservation numérique, citons iPRES et PASIG ; - La *Digital Preservation Coalition* organise des journées d'information sur des sujets particuliers liés à la préservation numérique.

Ressources

A Preservation Primer
http://knconsultants.org/a-preservation-primer/
Ce petit guide pratique et clair sur la préservation, destiné aux débutants, a été rédigé par le personnel de Portico. Il résume les problèmes et décrit les différentes options de préservation à court et à long terme qu'une organisation peut prendre pour planifier la préservation numérique à long terme de ses contenus, en commençant par la protection à court terme et en terminant par la préservation complète et la protection à long terme (83 pages).

Putting Parsimonious Preservation into Practice
http://www.nationalarchives.gov.uk/documents/information-management/parsimonious-preservation-in-practice.pdf
Le principe de la « préservation parcimonieuse » a été développé à l'origine en 2009 aux Archives nationales du Royaume-Uni comme une approche pour les institutions de petite ou moyenne taille afin de leur permettre de commencer à travailler sur la préservation numérique, mais il est également pratique pour les institutions de grande envergure. Ce principe détermine les conseils et les orientations donnés au secteur des archives du Royaume-Uni en matière de préservation numérique (11 pages).

Community Owned digital Preservation Tool Registry COPTR
http://coptr.digipres.org/Main_Page

[4] Cette liste a été fermée (NdT).
[5] Le blog a été redirigé vers l'adresse suivante : https://www.dpconline.org/knowledge-base/tags/122-digipresnews

Le registre COPTR décrit des outils utiles pour la préservation numérique à long terme et agit principalement comme un outil de recherche et d'évaluation pour aider les professionnels à trouver les outils dont ils ont besoin pour préserver les données numériques. Le registre COPTR vise à rassembler en un seul endroit les connaissances de la communauté de la préservation numérique sur les outils de préservation. Il a été initialement alimenté par des données provenant de registres gérés par les organisations partenaires du registre COPTR, notamment ceux tenus par le *Digital Curation Centre*, le *Digital Curation Exchange*, la *National Digital Stewardship Alliance*, l'*Open Preservation Foundation* et le *Preserving digital Objects With Restricted Resources project* (POWRR). Le registre COPTR rassemble des détails factuels de base sur un outil, ses fonctions, les moyens de trouver plus d'informations (URL pertinentes) et des références aux expériences des utilisateurs avec l'outil. Le champ d'application est une interprétation large du terme « préservation numérique ». En d'autres termes, si un outil est utile pour effectuer une fonction de préservation numérique telle que celles décrites dans le modèle OAIS ou le modèle de cycle de vie DCC, alors il entre dans le champ d'application de ce registre.

Ateliers de préservation numérique de la DPC

https://www.dpconline.org/digipres/train-your-staff/training-resources

Les ateliers de la DPC, « *Getting Started in Digital Preservation »,* sont des événements conçus pour sensibiliser aux questions de préservation numérique, accroître la participation aux activités de préservation numérique et afficher le soutien et les ressources disponibles pour vous aider dans votre travail. Ils fournissent une introduction à la préservation numérique, permettent de mieux comprendre les risques encourus par les objets numériques, comprennent des sessions pratiques pour vous aider à appliquer la planification et les outils de préservation numérique, et présentent des conférenciers qui partagent leur propre expérience de mise en pratique de la préservation numérique. Vous trouverez les détails des prochains ateliers, ainsi que les programmes et présentations des conférenciers des ateliers précédents sur la page des événements de la DPC.

Digital Preservation Management: Implementing Short-Term Strategies for Long-Term Solutions[6]

dpworkshop.org/

Un excellent tutoriel en ligne gratuit qui vous présente les principes de base de la préservation numérique. Il est particulièrement destiné aux bibliothécaires, archivistes, conservateurs, gestionnaires et spécialistes techniques. Il comprend des définitions, des concepts clés, des conseils pratiques, des exercices et des références actualisées. Le tutoriel est disponible en anglais, français et italien.

Canadian Heritage Information Network (CHIN) Digital Preservation Toolkithttps://www.canada.ca/en/heritage-information-network/services/digital-preservation/toolkit.html

Le RCIP a publié une série de documents pour identifier les objets numériques trouvés dans les musées, le risque potentiel et l'impact de la perte d'objets, et comment commencer à élaborer des politiques, des plans et des procédures de préservation. La boîte à outils comprend un modèle d'inventaire de préservation numérique, des directives pour l'élaboration d'un cadre de politique

[6]Gestion de la préservation numérique : mise en œuvre de stratégie à court terme pour des solutions à long terme

de préservation numérique, des arbres de décision et un modèle de plan de préservation numérique.

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

https://tpverso.files.wordpress.com/2016/03/swierczek.pptx

Cette présentation de Julie Swierczek, ancienne responsable des objets numériques et archiviste numérique aux musées d'art de l'université de Harvard, constitue une bonne sensibilisation et explication de la préservation numérique destinées à d'autres collègues institutionnels non spécialistes, notamment « le désespoir des archivistes face au grand public qui confond les archives avec les copies de sauvegarde des données » (la présentation contient 82 diapositives, mais de nombreuses images et notes de diapositives sont incluses, ce qui rend le contenu facilement compréhensible).

The National Archives Digital Continuity Guidance [7] http://www.nationalarchives.gov.uk/information-management/manage-information/policy-process/digital-continuity/step-by-step-guidance/step-2/

Ces conseils vous guident tout au long du processus de création d'un registre des biens informationnels et comprennent un modèle au format de feuille de calcul Excel. Le registre peut être utile aux gestionnaires de documents/informations en tant que modèle dont ils peuvent démontrer l'adéquation avec la gestion des risques de l'entreprise.

Risk Management for Digital Preservation

https://vimeo.com/171082277

Cette vidéo, tirée d'une série de vidéos couvrant les sujets des tournées de la présentation « Getting Started in Digital Preservation », fournit une brève introduction à l'utilisation de la gestion des risques pour la préservation numérique.

[7] Guide de continuité numériques des Archives nationales du Royaume-Uni

Études de cas

Étude de cas de la bibliothèque *Bishopsgate*
http://wiki.dpconline.org/index.php?title=Bishopsgate_library_case_study
Un audit des collections et une analyse de rentabilité ont constitué les premières étapes de la préservation numérique à la bibliothèque de l'Institut *Bishopsgate* (28 pages).

Starting Small: Practical First Steps in Digital Preservation
http://www.slideshare.net/hakbailey/starting-small-practical-first-steps-in-digital-preservation-13385434
Un exemple de la manière dont les principes de préservation numérique peuvent être ajoutés aux activités de gestion des collections d'une petite institution (*Dartmouth College* USA de 2010 à 2012), sans nécessiter beaucoup de ressources supplémentaires (26 diapositives).

Note d'étude de cas de la DPC : le *West Yorkshire Archive Service* accepte une collection numérique
http://www.dpconline.org/component/docman/doc_download/511-casenotemlawyas.pdf
Dans cette étude de cas financée par le Jisc, le personnel du service des archives du *West Yorkshire* rend compte de son expérience en matière de collecte de leurs premières grandes archives numériques. Ils évoquent les nouveaux problèmes et les nouvelles méthodes de travail auxquels ils ont été confrontés. Ils concluent que « Si nous essayons, nous pouvons échouer ; si nous n'essayons pas, nous échouerons certainement ». Octobre 2010 (4 pages).

Note d'étude de cas de la DPC : le musée de Glasgow fait les premiers pas pour transformer un casse-tête d'histoire orale en une opportunité
http://www.dpconline.org/component/docman/doc_download/502-casenoteglasgowmuseums.pdf
Cette étude de cas financée par le Jisc examine comment les musées de Glasgow ont pris quelques mesures simples pour aborder la préservation numérique et ont créé des opportunités à court et long terme. Des activités telles que la création d'un inventaire, l'évaluation de l'importance et la promotion de l'accès constituent la base de la confiance nécessaire pour gérer les défis plus larges, et peuvent apporter des récompenses rapides si elles sont correctement intégrées dans la mission d'une organisation. Septembre 2010 (4 pages).

Étude de cas sur la planification de la préservation numérique
http://www.dpconline.org/component/docman/doc_download/863-2013-may-getting-started-london-planning-case-study-ed-fay
Ensemble de diapositives provenant de la présentation de l'atelier DPC *Getting Started in Digital Preservation* par Ed Fay publié en mai 2013. Il s'agit d'un excellent aperçu concis de la planification de la préservation numérique et de la manière de l'aborder (20 diapositives).

Références

Dollar, C.M. and Ashley, L.J., 2014.*Assessing Digital Preservation Capability Using a Maturity Model Process Improvement Approach*. Disponible à l'adresse suivante :

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. Disponible à l'adresse suivante : http://www.digitalpreservation.gov/ndsa/working_groups/documents/NDSA_Levels_Archiving_2013.pdf

Version française : Thomas Ledoux, Jordan de la Houssaye, Stéphane Reecht, Bertrand Caron, Megan Phillips, et al. *NDSA Levels of Preservation* version 1, traduction française. 2019, disponible à l'adresse suivante : https://hal-bnf.archives-ouvertes.fr/hal-02162334/document

Stratégies institutionnelles

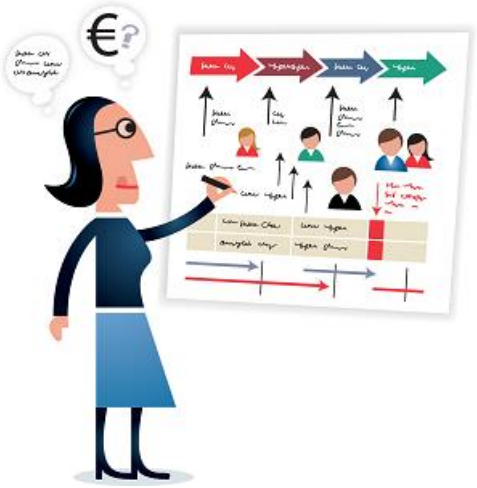


Image: digitalbevaring.dk CC BY 2.5 Denmark

A qui s'adresse-t-on ?

Cette section s'adresse tant aux administrateurs généraux (DigCurV: *Executive Lens*) qu'aux encadrants (DigCurV: *Manager Lens*) au sein des institutions. Les prestataires de services tiers existants ou potentiels sont également concernés.

Niveau de connaissance présumé

Intermédiaire (compréhension de base des enjeux, un peu d'expérience pratique).

Objectif

- Former la base des développements ultérieurs de politiques et de stratégies adaptées aux différentes institutions ;
- Fournir des exemples de bonnes pratiques qui pourraient servir de modèles ;
- Cette section décrit un certain nombre de stratégies utilisées avec succès par les institutions pour développer des approches de la préservation numérique. Chaque sous-section traite de l'approche, de ses avantages et inconvénients potentiels, puis fournit des exemples de l'approche ainsi que des lectures complémentaires sur le sujet. De telles stratégies constitueront un élément essentiel de l'élaboration de la politique d'entreprise pour aborder la préservation numérique. L'élaboration d'une politique solide combinée à des pratiques et des procédures de travail efficaces (voir Activités organisationnelles), a été essentielle à l'efficacité des programmes de préservation numérique.

Politiques et stratégies institutionnelles

Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Introduction

L'objectif de cette section est d'aider les institutions à comprendre, développer et mettre en œuvre des politiques et stratégies de préservation numérique. Celles-ci doivent permettre aux institutions de fixer des objectifs de préservation numérique, des priorités et des mécanismes qui soutiendront également la collecte, la gestion du cycle de vie et la diffusion des objets numériques.

Politique et stratégie sont des termes qui sont souvent utilisés de manière interchangeable ou dans une logique hiérarchique variable selon les institutions. Par souci de cohérence, le présent manuel définit la « politique » comme le document de niveau le plus élevé et la « stratégie » comme les documents et les procédures qui déclinent la mise en œuvre de la politique. En principe, l'élaboration d'une politique précède l'élaboration d'une stratégie. À son tour, la stratégie peut être développée ou révisée/revue régulièrement, alors que la politique peut avoir un cycle de révision plus long. Ainsi, une politique sert les objectifs généraux de l'organisation tandis que les stratégies individuelles peuvent servir ses différentes entités.

Les documents de politique et de stratégie constituent une base sur laquelle toutes les activités relatives à la gestion des documents numériques peuvent être appuyées. Des documents de politique et de stratégie bien conçus et élaborés de façon collaborative sont la garantie d'un niveau élevé d'appropriation et l'assurance que les activités quotidiennes de gestion des documents numériques s'y conforment réellement. En retour, cela donne la certitude que les objets numériques sont gérés de manière appropriée et optimale. Les documents de politique constituent également la base de la planification des coûts et des demandes de financement. La stratégie peut être utilisée comme un outil flexible pour s'adapter aux changements de situation et démontrer que des enseignements ont été tirés des retours d'expérience.

Dans toute institution, il y a un ensemble d'acteurs pour lesquels la gestion du cycle de vie des objets numériques est un enjeu. Ces acteurs peuvent contribuer à la gestion de ces objets numériques, ils peuvent créer ou gérer des métadonnées associées à ces objets, ou ils peuvent avoir la responsabilité de la gestion de collections. Les utilisateurs finaux sont également des acteurs clés, car leurs besoins déterminent ce qui est important pour la préservation. Les points de

vue des différents acteurs et leurs rôles par rapport à la gestion des objets numériques doivent être envisagés tant au niveau de la politique que de la stratégie.

Il pourra être utile d'appliquer la méthode de gestion itérative en quatre étapes *Plan-Do-Check-Act* (planifier-réaliser-vérifier-agir) comme un modèle d'amélioration continue, de développement, de mise en œuvre et de révision des politiques et des stratégies.

Politique de préservation numérique dans un contexte organisationnel plus large

Si vous vous lancez ou envisagez de vous lancer dans l'élaboration d'une politique de préservation numérique pour votre organisation, il est nécessaire de commencer par analyser le contexte dans lequel cette politique existera.

Il est probable qu'un large éventail de documents de politiques existe déjà au sein de votre organisation, couvrant une grande variété de sujets tels que la dotation en personnel, les technologies de l'information, l'évaluation des risques et les finances. Il peut également exister un certain nombre de politiques relatives à des sujets plus spécifiques de gestion des documents et des collections, possiblement pertinentes pour les activités de préservation numérique. Il est essentiel de tenir compte à la fois du contenu, du style et de la structure habituelle des politiques pertinentes au sein de l'organisation ainsi que de la manière dont la politique de préservation numérique s'inscrira dans un paysage plus large. Aucune politique ou stratégie ne peut se suffire à elle-même : pour obtenir du soutien et réussir la mise en œuvre de la politique de préservation, il est essentiel de l'inscrire dans un contexte institutionnel plus large.

Considérations importantes dans l'élaboration de la politique et de la stratégie

Un aspect important de l'élaboration de la politique est la prise en compte des besoins spécifiques de votre organisation et de ses principaux pilotes. Aligner l'activité de l'organisation sur les pilotes garantit que la stratégie et sa mise en œuvre soient également alignées sur les besoins de l'activité. Les documents de politique et de stratégie devraient s'articuler explicitement, avec, les politiques et stratégies préexistantes pertinentes et s'appuyer sur les pratiques en vigueur. La collaboration, le partage et la consultation avec les acteurs sont des éléments essentiels dans l'élaboration de la politique et de la stratégie.

Les politiques de préservation numérique sont idéalement neutres sur le plan technologique, c'est-à-dire qu'elles ne dépendent pas d'une plate-forme ou d'un système technologique particulier. Toutefois, dans la réalité, cela peut s'avérer impossible. Dans ce cas, elles devraient être axées sur des principes, des buts et des objectifs que la technologie disponible peut soutenir.

Afin de développer une documentation et des processus clairs, cohérents et solides, il est essentiel d'adhérer à une méthodologie définie et d'établir un plan de révision afin que la politique et la stratégie restent pertinentes et à jour.

1. **Établir l'objectif**. La première étape consiste à définir l'objet principal de la politique de préservation numérique, son périmètre et ses principaux objectifs. Ceux-ci permettront de maintenir le processus d'élaboration de la politique ciblé et son contenu cohérent. À ce stade, il convient de réfléchir à la manière dont le document sera utilisé, à la fois comme un outil de sensibilisation et une aide pour orienter la création et la mise en œuvre de la stratégie.

2. **Rechercher**. Comme indiqué ci-dessus, il est essentiel de comprendre le contexte organisationnel dans lequel la politique existera. Il conviendra de consacrer du temps à l'étude

de la politique existante, à la compréhension des pilotes de l'activité dans l'organisation et aux besoins des principaux groupes d'acteurs. Cette phase comprendra également des recherches sur les meilleures pratiques en matière de politique et de stratégies de préservation numérique, un examen des outils et des ressources disponibles ainsi qu'une étude des politiques et des stratégies d'autres organisations. De nombreuses ressources sont disponibles, avec des suggestions sur ce qu'il faut inclure dans votre politique et votre stratégie de préservation numérique (voir la rubrique Ressources).

3. **Identifier les éléments et développer la structure**. Sur la base des recherches effectuées lors de la phase précédente, il conviendra de sélectionner les principaux sujets et questions à aborder dans la politique et la stratégie. Il est essentiel d'élaborer une structure claire pour les documents afin de garantir leur utilité pratique et de faciliter les mises à jour et les révisions. La structure devra prendre en compte toute norme ou bonne pratique existante de rédaction des documents de politique et de stratégie au sein de l'organisation.

4. **Développer le contenu**. Le contenu de la politique doit être de haut niveau et fixer des buts et des objectifs larges. Il doit éviter au maximum la mention d'éléments spécifiques tels que les détails liés à des solutions technologiques particulières, bien qu'il puisse contenir une référence à des engagements établis au niveau de l'organisation. Les informations sur l'application pratique de la politique seront définies par les documents stratégiques pertinents. Le contenu peut également être ambitieux par rapport aux buts et aux objectifs, mais il faut veiller à ne pas fixer des buts impossibles à atteindre. Des recommandations sur la manière d'aborder des questions spécifiques dans le cadre de votre politique et de votre stratégie sont disponibles au sein de nombreuses ressources (voir Ressources).

5. **Interroger les acteurs**. Il est essentiel d'obtenir l'adhésion des différents groupes d'acteurs pour s'assurer que votre politique et votre stratégie sont adaptées à leurs objectifs et bénéficieront de soutien au sein de l'organisation. La présentation des projets de documents aux principaux groupes d'acteurs est une partie importante du processus de rédaction et tout retour d'information fourni doit être soigneusement pris en compte. Cela peut également constituer une étape clé dans la sensibilisation à la préservation numérique au sein de votre organisation, en permettant aux acteurs de se sentir impliqués dans le processus et de comprendre comment les activités de préservation numérique sont reliées à leur propre travail (voir Sensibilisation).

6. **Obtenir l'approbation**. La plupart des organisations exigeront que les nouveaux documents de politique soient officiellement ratifiés par la direction. Assurez-vous de bien connaître le processus de l'organisation et les exigences qui devront être remplies. Une fois ratifiée, la politique aura plus de poids et sera donc plus facile à mettre en œuvre dans le cadre d'une stratégie continue.

7. **Réviser régulièrement**. Les documents de politique et de stratégie ne doivent pas être statiques et doivent répondre à l'évolution des besoins des acteurs, du contexte organisationnel général et des mises à jour des bonnes pratiques. Un cycle de révision régulière doit être établi mais peut également être induit par des changements significatifs dans l'un des domaines mentionnés ci-dessus.

8. **Mettre en œuvre**. Établissez un plan de mise en œuvre pour concrétiser la politique et la stratégie dans vos activités quotidiennes. N'oubliez pas qu'elles sont un moyen pour atteindre une fin et non une fin en soi.

Ressources

Digital Preservation Policies Study
http://www.webarchive.org.uk/wayback/archive/20140615022334/http://www.jisc.ac.uk/media/documents/programmes/preservation/jiscpolicy_p1finalreport.pdf
Cette étude financée par le Jisc et publiée en 2008 a créé un modèle de politique de préservation numérique et les modalités de mise en œuvre qui l'accompagnent, en se basant sur l'examen des politiques de préservation numérique existantes. Bien qu'elle se concentre sur les secteurs de l'enseignement supérieur et de la formation continue au Royaume-Uni, l'étude s'inspire largement des politiques et des réalisations d'autres secteurs et pays.

En outre, l'étude comporte une série de cartographies des liens de la préservation numérique avec d'autres stratégies institutionnelles clés dans les universités et les collèges britanniques, dans le but d'aider les institutions et leur personnel à développer des politiques de préservation numérique appropriées et des modalités de mise en œuvre établies dans le contexte de stratégies institutionnelles plus larges (60 pages).

Digital Preservation Policies: Guidance for Archives
http://www.nationalarchives.gov.uk/documents/information-management/digital-preservation-policies-guidance-draft-v4.2.pdf
Ce guide publié par les Archives nationales du Royaume-Uni en 2011 explique les principales caractéristiques d'une politique de préservation numérique. Il explique pourquoi une politique est nécessaire et comment elle soutient la préservation numérique. Ce guide s'adresse principalement aux services publics d'archives (16 pages).

Analysis of Current Digital Preservation Policies
http://www.digitalpreservation.gov/documents/Analysis%20of%20Current%20Digital%20Preservation%20Policies.pdf?loclr=blogsig
Ce rapport publié en 2013 par Madeline Sheldon, et le NDIIPP à la Bibliothèque du Congrès, traite de l'état actuel de la planification de la politique de préservation numérique au sein des organisations du patrimoine culturel. L'ensemble des politiques ou stratégies de préservation numérique nouvelles ou récemment révisées, publiées de 2008 à 2013, a donné lieu à une analyse de haut niveau du contenu de ces documents. Un résumé des conclusions a également été publié sur le blog *The Signal* (23 pages).

APARSEN D35.1 Exemples de structures de bonne gouvernance et de politiques de données
http://www.alliancepermanentaccess.org/index.php/consultancy/member-resources/documents-and-downloads/?did=174[8]

[8] NdT : Le document n'était plus accessible en ligne à la date de la présente traduction.

Ce rapport résume le niveau d'adhésion à des politiques de gouvernance et de données interopérables. Il se termine par une sélection de recommandations qui devraient être prises en compte lors de l'élaboration de politiques de données concernant la préservation numérique (2014, 43 pages).

Catalog of Preservation Policy Elements
http://wiki.opf-labs.org/display/SP/Catalogue+of+Preservation+Policy+Elements
Le projet européen SCAPE (2011-2014) a été chargé d'examiner les politiques existantes et de produire un catalogue d'éléments de politique pour aider ceux qui rédigent ces documents. Ce wiki donne des informations sur le contexte de ce travail et contient des pages pour chaque élément que le projet SCAPE a suggéré aux organisations de prendre en compte lors de la rédaction de politiques, en mettant l'accent sur les activités de planification et de surveillance. Le rapport final[9] de ce travail a également été rendu public en février 2014 sur le site web SCAPE.

Published Preservation Policies
http://wiki.opf-labs.org/display/SP/Published+Preservation+Policies
Un vaste répertoire web préparé par le projet SCAPE en 2015 énumérant les politiques de préservation numérique qui sont publiquement disponibles, en ligne, pour les bibliothèques, les archives, les centres de données et diverses institutions.

Études de cas

A Digital Preservation Policy for Parliament
http://www.parliament.uk/documents/upload/digitalpreservationpolicy1.0.pdf
L'objectif de cette politique publiée en 2009 est d'établir et de communiquer les principes qui guident les activités du Parlement britannique pour assurer la préservation de ses informations numériques. D'autres documents de politique, procédures, normes et orientations seront élaborés à l'avenir pour traiter des aspects stratégiques spécifiques (17 pages).

Politique de préservation numérique de *Hampshire Archives and Local Studies* (HALS)
http://documents.hants.gov.uk/archives/HRODigitalPreservationPolicy.pdf
Pour faire face au risque de perte des objets numériques, HALS a élaboré une politique et une stratégie de préservation numérique. La politique décrit l'approche du *Record Office* en matière de préservation numérique, tandis que l'objectif de la stratégie est de décrire cette approche de manière plus détaillée, incluant les spécifications techniques le cas échéant.

Note d'étude de cas de la DPC – Les papiers du Cabinet : la politique comme mesure de l'engagement
http://www.dpconline.org/component/docman/doc_download/449-casenotecabinetpapers.pdf
Cette étude de cas des Archives nationales du Royaume-Uni examine la relation entre la politique et la pratique en matière de préservation numérique. Les organismes financeurs devraient demander des copies des politiques de préservation numérique du candidat lorsqu'ils financent la création de données, car celles-ci sont une indication de l'engagement de l'organisme en matière d'accès à long terme. Les Archives nationales ont numérisé une part importante des documents du

[9] NdT : Le document n'était plus accessible en ligne à la date de la présente traduction, une archive est disponible sur https://web.archive.org/web/20201229173254/https://scape-project.eu/category/deliverable.

Cabinet britannique et disposent d'un cadre politique soigneusement étudié pour la gestion à long terme des ressources numériques. Mai 2010 (3 pages).

Note d'étude de cas de la DPC - Les revues galloises en ligne : un leadership efficace pour un objectif commun

http://www.dpconline.org/component/docman/doc_download/450-casenotewelshjournals.pdf

Cette étude de cas financée par le Jisc examine un projet de numérisation complexe à la *National Library of Wales*, un exemple d'organisation où les acteurs sont nombreux et où de nombreuses compétences différentes sont requises. Le fait de nommer un seul membre du personnel de haut niveau comme responsable de la préservation numérique et de lui permettre de travailler dans différentes sections de l'institution a atténué le risque d'incertitude quant à la responsabilité des actions de préservation. Juin 2010 (3 pages).

Stratégie de préservation numérique de la British Library 2013-2016

La stratégie de la *British Library* comprend quatre priorités. Chaque priorité est accompagnée d'une série d'actions. Ces priorités sont alignées sur l'approche globale de la Bibliothèque en matière de préservation des collections et sur ses cinq principes de gestion durable : prévoir, protéger, prioriser, préserver et développer.

Priorité stratégique n°1 : veiller à ce que notre dépôt numérique puisse stocker et préserver nos collections sur le long terme.
Priorité stratégique n°2 : gérer les risques et les défis associés à la préservation numérique tout au long du cycle de vie du contenu des collections numériques.
Priorité stratégique n°3 : intégrer la pérennité numérique comme principe organisationnel pour la planification et le développement des bibliothèques numériques.
Priorité stratégique n°4 : tirer parti de la collaboration avec d'autres institutions nationales et internationales sur les initiatives de préservation numérique.

La version complète en PDF de la stratégie contient plus de détails sur chacune de ces priorités (16 pages).

La stratégie disponible sur le site de la *British Library* est celle de 2017 à 2020 :
https://www.bl.uk/digital-preservation/strategy.

Politique de préservation de la bibliothèque *Wellcome*

http://wellcomelibrary.org/what-we-do/library-strategy-and-policy/preservation-policy/

L'objectif de la politique de préservation de la bibliothèque *Wellcome* est de fournir une déclaration complète sur la préservation et la conservation des collections de la bibliothèque. Elle prend en compte tous les objets dans tous les formats. La politique comprend trois parties qui couvrent les déclarations générales, la gestion des documents physiques et la gestion des objets numériques (25 pages).

Politique de préservation de *UK Data Archive*

http://data-archive.ac.uk/media/54776/ukda062-dps-preservationpolicy.pdf

Cette politique publiée en 2014 expose les principes sur lesquels reposent les principales activités de *UK Data Archive* (l'Archive) : la préservation active des ressources numériques pour l'utilisation et la réutilisation au sein de sa communauté principale d'utilisateurs. Du point de vue de la préservation, cette politique est globalement conforme au modèle de référence OAIS, avec des ajouts et des modifications qui sont spécifiques aux objets conservés par l'Archive. L'Archive a une

série d'exigences strictes pour ses activités de préservation numérique. Ces exigences sont définies dans la politique. La manière dont ces exigences peuvent être atteintes au mieux au regard des exigences réglementaires, des bonnes pratiques archivistiques, de la sécurité de l'information et des fonds disponibles est également détaillée. En conséquence, la politique de préservation des archives est basée sur des formats de fichiers ouverts et disponibles, la migration des données et le rafraîchissement des supports (16 pages).

Digital Preservation Strategies for a Small Private College

http://files.archivists.org/pubs/CampusCaseStudies/CASE-16-MegMiner-Final.pdf

Le projet POWRR (2011 - 2014) a étudié, évalué et recommandé des solutions de préservation numérique évolutives pour les bibliothèques disposant de petites quantités de données et/ou de peu de ressources. Les « bonnes pratiques » établies en matière de préservation numérique ne permettent pas de faire face aux réalités quotidiennes des services qui ne peuvent pas consacrer de moyens ou de personnel aux processus de préservation. Meg Miner, de l'Université Weslyan de l'Illinois, explique ce qui peut être fait pour assurer une bonne gestion des documents institutionnels numériques et numérisés avant qu'un système de préservation complet ne soit mis en place. 2015 (13 pages).

Developing a Digital Preservation Policy

http://www.dpconline.org/component/docman/doc_download/1321-making-progress-hsbc-nov-2014-lee

Une excellente présentation et étude de cas par Kirsty Lee de l'Université d'Édimbourg à l'atelier DPC *Making Progress in Digital Preservation,* en octobre 2014, expliquant la méthodologie qu'elle utilise pour construire une politique de préservation numérique à Édimbourg (14 pages).

Collaboration

Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Introduction

Il existe des raisons impérieuses et, dans certains cas, des injonctions politiques, pour s'engager dans une plus grande collaboration au sein des organisations et entre elles afin d'affronter et de surmonter efficacement les défis de la préservation numérique. L'éventail des compétences requises exige une certaine souplesse au sein des structures organisationnelles afin de faciliter le travail en équipes pluridisciplinaires. Les questions de préservation numérique auxquelles sont confrontées toutes les organisations et tous les secteurs se recoupent largement. Il est donc logique de mettre en commun les compétences et l'expérience. La communication avec les principales parties prenantes, en utilisant des termes et un langage qu'elles comprennent (voir la section Sensibilisation) jouera un rôle majeur dans l'établissement et le maintien des collaborations.

Collaboration interne

L'hypothèse habituelle est que la collaboration est externe. Cependant, la plupart des bibliothèques et archives géreront une combinaison de ressources papier et numériques dans un avenir proche et devront structurer leur organisation pour gérer les besoins disparates des deux. L'effacement des frontières et les changements de cycle de vie que la technologie numérique provoque signifient que des équipes structurellement distinctes devront désormais coopérer afin de préserver et gérer de manière globale des objets, qu'ils soient numériques ou non.

Une telle coopération et un tel travail en commun pourraient bien s'avérer impossibles à moins que des mécanismes ne soient mis en place pour les faciliter et qu'il n'y ait une adhésion et un parrainage clairs de la part des dirigeants pour promouvoir ces actions. Au niveau stratégique, un comité interdisciplinaire ou une équipe projet chargée d'élaborer et de superviser les objectifs est une manière de s'assurer que toutes les équipes concernées peuvent être réunies. Au niveau opérationnel, il faudra envisager de définir les tâches spécifiques requises et répartir de manière logique les responsabilités. La mise en place de groupes de travail chargés d'étudier des questions spécifiques est un moyen de combiner l'éventail des compétences requises. Une bonne communication et sensibilisation à destination des autres parties prenantes seront également importantes (voir Sensibilisation).

Si la nécessité d'une politique ou d'une stratégie relative à la préservation numérique peut être bien établie au sein de l'équipe de l'entrepôt numérique, le moteur de la collaboration interne est généralement la réponse à un défi spécifique auquel une organisation est confrontée.

Avantages

- Fait bon usage des compétences et de l'expertise disponibles et fait valoir que la préservation numérique est une question institutionnelle et non une question qui relève exclusivement de l'entrepôt numérique.
- Favorise le travail en équipe en améliorant la compréhension des objectifs communs et l'identification des personnes qui doivent y contribuer.
- Plus tôt la préservation numérique fera partie du travail quotidien d'une organisation et de ses employés, plus la préparation et la transition vers le monde numérique seront facilitées.
- Reconnaît la diversité des compétences requises dans l'environnement numérique en général et la préservation numérique en particulier.
- Est plus susceptible de se concentrer et de s'aligner sur les objectifs et les priorités de l'institution.

Conserve un haut niveau de compétence pour ce travail.

Inconvénients

- Cela peut être frustrant et prendre du temps à court terme.
- La communication peut être difficile au début – voir par exemple, les problèmes de terminologie avec le terme « archives » qui a une signification différente pour les archivistes et les collègues des DSI.
- Donner l'impression de perdre le contrôle peut être perçu comme un risque pour la direction générale.
- Le personnel peut se sentir mal à l'aise avec les nouvelles méthodes de travail.
- Les structures organisationnelles peuvent ne pas être suffisamment souples pour faciliter une collaboration efficace entre les différentes sections.

Collaboration externe

Il peut y avoir un certain nombre de motivations à développer des collaborations externes. Il y a le simple désir pour les spécialistes isolés de travailler avec d'autres collègues professionnels et de rechercher une validation externe de leurs idées ou de leur orientation. À l'autre bout de l'échelle, il y a la réponse aux possibilités de financement externe, les financeurs mettant désormais davantage l'accent sur la collaboration. Quelques exemples de collaboration externe dans le secteur de la préservation numérique sont présentés ci-dessous :

- **Collaboration autour d'un problème spécifique pour progresser plus facilement et de manière plus abordable**. *Digital Preservation Coalition* elle-même en est un exemple au Royaume-Uni. Les membres sont encouragés à s'engager et à collaborer sur un certain nombre de questions différentes liées à la préservation numérique, tant à un niveau général que sur des sujets spécifiques. Un autre exemple est la *Section for Archives and Technology* de l'*Archives and Records Association*, qui réunit les membres de la profession pour examiner des aspects spécifiques du travail et partager les pratiques actuelles.

- **Collaboration autour d'une norme. On peut citer comme exemple l'appel à collaboration** lancé en 2015 autour de la révision du modèle de référence OAIS. Dans le cadre d'une

initiative coordonnée par la DPC, les professionnels travaillant dans ce domaine ont été encouragés à s'engager et à contribuer à une réponse commune. (Voir http://wiki.dpconline.org/index.php?title=OAIS_Community)

- **Collaboration autour d'un logiciel ou d'un système spécifique**. On peut citer comme exemple les groupes d'utilisateurs qui travaillent à l'élaboration de solutions logicielles de préservation numérique, à la fois commerciales et libres. Lorsque l'on examine un logiciel pour la première fois, il est très important de pouvoir partager son expérience et ses connaissances.

- **Collaboration au sein d'une zone géographique spécifique**. Il existe de nombreux exemples d'organisations qui collaborent en raison de leur proximité géographique et de la facilité de travail en commun qu'elle offre. Le *Digital Preservation Group* de l'ARCW (*Archives & Records Council Wales*) en est un exemple (voir Études de cas).

Avantages

- Une organisation engagée avec une mission clairement définie.
- Formaliser un accord permet une répartition claire des responsabilités entre les partenaires.
- Une meilleure compréhension des questions complexes.
- Un plus grand avantage pratique de la mise en commun des ressources et de l'expertise.
- Une meilleure réputation grâce à la réussite d'un projet ou en montrant qu'on est capable de gérer la préservation numérique.
- Une amélioration des perspectives pour une collaboration future.

Inconvénients

- Difficulté d'établir des accords non ambigus pouvant être acceptés par toutes les parties.
- Temps nécessaire pour mettre en place des équipes ou un cadre de collaboration.
- Difficultés de communication entre des personnes appartenant à des contextes professionnels et organisationnels différents.
- Obstacles bureaucratiques potentiels.

La collaboration externe peut être informelle ou formelle. Le partage d'expérience et la collaboration informelle font souvent partie de la formation professionnelle continue d'un individu. Les collaborations plus importantes et plus complexes sont davantage susceptibles de nécessiter la conclusion d'un accord de partenariat formel qui peut être utile pour définir la portée et les limites de la relation de travail et répartir spécifiquement les responsabilités.

Ressources

Benefits from Research Data Management in Universities for Industry and Not-for-Profit Research Partners
http://opus.bath.ac.uk/32509/
Application d'une cartographie des parties prenantes en utilisant le cadre des avantages du KRDS pour examiner les avantages de la gestion des données associée à des collaborations de recherche entre l'Université de Bath et l'industrie et des organismes à but non lucratif . Ce document présente une liste récapitulative des avantages pour les différentes parties prenantes qui peuvent découler de ces collaborations en termes de gestion et de préservation des données de recherche.

Aligning National Approaches to Digital Preservation Conference
(Actes de la conférence de 2012)
http://educopia.org/publications/anadp
Cette publication contient une collection d'essais évalués par des pairs, qui ont été élaborés par des groupes d'experts et des participants à la conférence. Elle vise à établir un ensemble de points de départ pour une meilleure harmonisation des initiatives de préservation numérique et souligne la nécessité de collaborations internationales stratégiques pour soutenir la préservation de notre mémoire culturelle collective (342 pages).

North West Region Digital Preservation Group
https://nwrdpg.wordpress.com/
Le *North West Region Digital Preservation Group* est un exemple de collaboration géographique informelle impliquant les autorités locales, les universitaires et les archivistes spécialisés. Les résultats comprennent des lignes directrices pour les déposants, un manuel pour les archivistes et des études pilotes sur l'archivage du web et du courrier électronique.

Études de cas

Note d'étude de cas–de la DPC - Les revues galloises en ligne : un leadership efficace pour un objectif commun
http://www.dpconline.org/component/docman/doc_download/450-casenotewelshjournals.pdf
Cette étude de cas financée par le Jisc examine un projet de numérisation complexe à la *National Library of Wales*, un exemple d'organisation où les parties prenantes sont nombreuses et où de nombreuses compétences différentes sont requises. Le fait de nommer un seul membre du personnel de haut niveau comme responsable de la préservation numérique et de lui permettre de travailler dans différentes sections de l'institution a atténué le risque d'incertitude quant à la responsabilité des actions de préservation. Juin 2010 (3 pages).

Note d'étude de cas–de la DPC - *Freeze Frame preservation partnerships*

http://www.dpconline.org/component/docman/doc_download/434-casenotefreezeframe.pdf

Cette note d'étude de cas examine la relation entre le projet *Freeze Frame*, d'une durée relativement courte, l'Institut de recherche polaire Scott et l'entrepôt numérique institutionnel qui a proposé de fournir des services de préservation à long terme pour garantir un accès continu à la fin du projet. L'étude montre que les petites organisations n'ont pas nécessairement besoin de mettre en place une infrastructure de préservation sophistiquée lorsqu'elles se lancent dans la numérisation. Le partenariat peut apporter des avantages inattendus aux deux parties, mais il doit être géré et documenté de manière réfléchie. Avril 2010 (4 pages).

Action communautaire via l'alliance britannique LOCKSS

http://www.slideshare.net/edinadocumentationofficer/ukla-dpc-final

Présentation donnée par Adam Rusbridge à la *Digital Preservation Coalition à Getting Started in Digital Preservation*, le 28 février 2011. Elle traite du rôle de l'alliance britannique LOCKSS et de la collaboration en matière de préservation des périodiques électroniques.

Archives & Records Council Wales Digital Preservation Working Group

http://www.nationalarchives.gov.uk/documents/Cloud-Storage-casestudy_Wales_2015.pdf

Cette étude de cas des Archives nationales du Royaume-Uni présente l'expérience d'un groupe de travail intersectoriel des archives galloises qui a coopéré pour tester une série de systèmes et de déploiements de services dans le cadre d'une validation de principe pour l'archivage dans le *cloud*. Elle explique le contexte organisationnel, la nature variée de leurs exigences et approches en matière de préservation numérique, ainsi que leur expérience en matière de sélection, de déploiement et de test de la préservation numérique dans le *cloud*. Janvier 2015 (10 pages).

Une infrastructure de collaboration pour un accès permanent au patrimoine numérique aux Pays-Bas

http://www.ncdd.nl/wp-content/uploads/2016/03/Network_Digital_Heritage_Netherlands.pdf[10]

En 2014, le *Network Digital Heritage* (NDE) a été mis en place par un groupe d'organisations nationales aux Pays-Bas. Le réseau a présenté une stratégie pour le développement d'une infrastructure nationale et interdisciplinaire de services pour le patrimoine numérique. L'un des programmes se concentre sur la préservation numérique (*Sustainable digital heritage*). L'objectif de ce programme est de travailler sur le partage, l'utilisation et la mise à l'échelle intersectorielle des services pour une préservation et un accès durables, tout en prêtant attention à la gestion des coûts et à la répartition des tâches. Ce programme est mis en œuvre par la NCDD, la *National Coalition for Digital Preservation* (3 pages).

Le projet SPRUCE

http://wiki.opf-labs.org/display/SPR/Home

Le projet SPRUCE (*Sustainable PReservation Using Community Engagement*) (2011-2013) a cherché à inspirer, guider, soutenir et permettre à l'enseignement supérieur et de recherche et aux institutions culturelles de combler les lacunes en matière de préservation numérique et d'utiliser les connaissances acquises dans le cadre de cette activité pour formuler un argument commercial convaincant en faveur de la préservation numérique. Cette collaboration multi-institutionnelle a rassemblé des archivistes et des experts en technologie par le biais d'événements de mashups et

[10] NdT : Le document n'était plus accessible en ligne à la date de la présente traduction

d'un hackathon. Les deux principaux résultats du projet ont été le *Business Case Toolkit* (https://wiki.dpconline.org/index.php?title=Digital_Preservation_Business_Case_Toolkit) et le COPTR (*Community Owned digital Preservation Tool Registry*) (voir https://coptr.digipres.org/index.php/Main_Page).

Combler le fossé de la préservation numérique : étude de cas

http://digital-archiving.blogspot.co.uk/2015/12/research-data-spring-case-study-for.html

Une collaboration entre les universités de Hull et de York. L'objectif du projet était de combler une lacune perçue dans les infrastructures de gestion des données de recherche existantes en ce qui concerne la préservation active des données. Les universités de Hull et de York disposaient toutes deux de dépôts numériques et d'une capacité de stockage suffisante, mais manquaient de systèmes et de processus pour assurer la préservation active des données.

Sensibilisation[11]

Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Introduction

La préservation numérique repose sur un large éventail de compétences et de services. Ceux qui en sont responsables doivent donc coordonner un ensemble diversifié de compétences, de politiques, d'outils et de services provenant de sources disparates. Pour certaines organisations, la préservation numérique est entièrement nouvelle et les ressources adéquates doivent être rassemblées pour la première fois. Les processus établis sont aussi confrontés à de nouveaux défis et, par conséquent, leur gamme d'outils et de services requis peut changer constamment. C'est pourquoi la capacité à communiquer avec d'autres membres du personnel, d'autres services et d'autres organisations est devenue une compétence clé pour les gestionnaires de la préservation numérique.

La technologie et le personnel ne cessant d'évoluer, la communication et la sensibilisation doivent être une préoccupation permanente plutôt que ponctuelle.

Lors des débuts de la préservation numérique, la communication et la sensibilisation impliquaient des déclarations abruptes sur les conséquences sociales et économiques de la perte et de l'obsolescence des données. Au fur et à mesure que des solutions ont émergé, les messages sont devenus plus subtils.

La sensibilisation consiste de plus en plus à identifier les parties prenantes et à les aider à comprendre :

- Comment leurs choix rendent les collections numériques plus ou moins sujettes à l'obsolescence ;
- Quels avantages leur apportera la gestion active d'objets numériques correctement formés et accessibles ;
- Pourquoi est-il nécessaire d'investir - qu'il s'agisse de temps, d'argent ou d'autres ressources - et quels investissements sont nécessaires pour bénéficier de ces avantages.

Dans un cadre institutionnel, cela signifie qu'il faut impliquer tous les agents concernés par le cycle de vie d'un objet numérique, les aider à établir des priorités et à soutenir les actions qui rendent et

[11] Digital Preservation Coalition, *Digital Preservation Handbook*, [en ligne] URL : https://www.dpconline.org/docs/digital-preservation-handbook2/1554-dp-handbook-institutional-strategies/file (lien consulté le 22 juin 2020)

maintiennent les collections pérennes, et décourager les actions qui mettent les collections en danger.

Analyse des parties prenantes

L'analyse des parties prenantes commence par une compréhension claire des objectifs de préservation numérique de l'organisation avant d'identifier les parties prenantes internes et externes qui peuvent influencer ces objectifs.

Après les avoir identifiés, il est possible d'élaborer un plan qui exposera vos objectifs et les intégrera au processus de préservation numérique. Une approche méthodologique claire permettra d'obtenir les meilleurs résultats et liera vos objectifs à un certain nombre d'autres activités de préservation numérique, telles que l'élaboration de politiques et de stratégies (voir Stratégies institutionnelles), la création d'une analyse de rentabilité (voir Analyses de rentabilité, avantages, coûts et impact) et l'identification des normes et des meilleures pratiques pertinentes (voir Normes et bonnes pratiques).

Les étapes suivantes contribueront à faciliter l'analyse approfondie des parties prenantes :

1 - Identifiez ce que vous souhaitez atteindre grâce à vos actions en matière de préservation numérique. La réflexion peut comprendre des listes des principales collections concernées et les principaux buts et objectifs ainsi que les bénéfices potentiels (voir les analyses de rentabilité, les bénéfices, les coûts et l'impact) qui en découleront. Cela constituera des références et objectifs clairs pour la défense de vos intérêts et pourra être adapté ultérieurement aux différents publics identifiés.

2 - Identifiez les groupes et les individus qui peuvent entraver ou permettre les activités de préservation numérique. Ceux-ci peuvent être internes ou externes, et une même partie prenante peut avoir plusieurs rôles. Par exemple, vous pouvez identifier le personnel des services informatiques comme un groupe clé qui peut ensuite inclure un responsable des services informatiques, des développeurs et du personnel chargé des fonctions support. Certains d'entre eux peuvent être facilement accessibles au sein de votre propre service, d'autres relèvent de différentes structures hiérarchiques, et d'autres encore sont entièrement externes. Cela signifie que vous devrez peut-être inclure d'autres managers ou responsables de services dans votre analyse des parties prenantes.

3 - Organisez les groupes de parties prenantes et les individus en publics clés qui sont en mesure d'influencer vos objectifs et vos priorités. Les publics choisis refléteront probablement les pratiques de travail de votre organisation et/ou votre approche de la préservation numérique. Peut-être ces publics seront-ils en rapport avec des parties spécifiques de votre structure organisationnelle (par exemple, la direction générale, les services informatiques, les gestionnaires de l'information) ou bien joueront-ils un rôle par rapport au processus de préservation numérique (par exemple, les financeurs, les déposants, –es utilisateurs).

4 - Établissez une étroite collaboration avec les principaux publics identifiés pour consolider les progrès vers les objectifs établis à l'étape 1. Comprendre les besoins, les priorités et les contraintes des parties prenantes internes et externes vous permettra d'obtenir des informations utiles à votre planification et améliorera votre compréhension des attendus des parties prenantes quant à la préservation numérique. Ces parties prenantes peuvent être contraintes par des limites budgétaires et/ou législatives dont il sera utile de prendre connaissance. Inversement, elles

peuvent également disposer d'une expertise ou de ressources pertinentes qui peuvent être déployées pour l'activité de préservation numérique. En outre, la compréhension du langage et de la terminologie utilisés par les parties prenantes améliore l'efficacité des stratégies de communication et peut contribuer à éviter les difficultés qui surviennent lorsque les parties prenantes interprètent mal un terme ou un concept. Votre capacité à utiliser le langage des parties prenantes aide généralement à faire adhérer collègues et collaborateurs à vos projets. Si les principales parties prenantes ont des intérêts conflictuels, vous devrez faire office de médiateur entre elles.

5 - Sur la base de cet engagement réciproque, définissez clairement les informations importantes à partager avec ces publics pour contribuer à les faire adhérer. Cela devrait comprendre :

- Des messages clés basés sur vos buts et objectifs. Il doit s'agir de déclarations simples et directes, rédigées dans un langage simple, afin qu'elles soient facilement comprises par un large éventail de publics non spécialisés. Idéalement, ces messages devraient également être alignés sur des stratégies et des objectifs organisationnels plus larges ;

- Les avantages que les parties prenantes retireront de leur participation ou de leur soutien aux activités de préservation numérique proposées. Un responsable informatique pourrait par exemple vouloir réduire les coûts de stockage en supprimant ou en dédoublonnant les éléments de stockage redondants. Une stratégie claire de préservation numérique peut les aider à réduire leurs besoins de stockage en distinguant les collections qui doivent être conservées de celles dont la préservation n'est plus nécessaire ;

- Ce qui sera exigé d'eux pour assurer le succès. Par exemple, vous souhaiterez peut-être définir des exigences claires en matière de métadonnées pour les déposants ; ou vous souhaiterez peut-être donner à votre service informatique des estimations concernant les quantités de stockage et la bande passante qui seront nécessaires et à quel moment ;

- Quels sont les obstacles et les idées fausses que vous devez aborder concernant la préservation numérique. La préservation est par exemple souvent confondue avec le stockage de copies de sauvegarde. Vous devrez peut-être adapter votre langage et votre terminologie à des publics spécifiques. Par exemple, certains termes comme « archivage » ont des significations différentes dans d'autres secteurs comme l'informatique.

6 – Élaborez un plan pour mobiliser chaque groupe de parties prenantes en vous basant sur votre connaissance de leurs priorités, de leur expertise et de leurs limites, et en utilisant les différents messages précédemment identifiés. Vous devrez peut-être utiliser des méthodologies différentes pour les différents groupes, en adaptant votre forme de communication pour qu'elle convienne au mieux au public et aux messages à transmettre. Il peut s'agir d'une série de canaux de communication comprenant des présentations, des documents d'information et des groupes de travail, ainsi que de l'élaboration de divers plans et ressources tels que des analyses de rentabilité (voir Analyses de rentabilité, avantages, coûts et impact), des politiques (voir Politiques et stratégies institutionnelles) et des registres des risques (voir Gestion des risques et du changement).

Préservation numérique dans les médias

La préservation numérique reçoit étonnamment peu d'attention dans les médias grand public. Les reportages sur le sujet ont tendance à tomber dans deux clichés : de lugubres histoires de perte de données et l'imminence d'une « ère noire numérique » ou des déclarations vues et revues sur le stockage indestructible.

La réalité est plus banale et plus subtile. Les exigences pratiques, détaillées et réalisables qui assurent un accès à long terme, telles que celles qui sont décrites dans ce manuel, attirent moins l'attention, mais peuvent apporter de réels avantages aux institutions et à leurs communautés d'utilisateurs.

Dans certains contextes de sensibilisation, il peut être utile de se référer à un vocabulaire courant pour expliquer des termes et concepts clés de la préservation numérique. Quelques exemples sont suggérés dans la section des ressources ci-dessous.

La communauté de la préservation numérique a créé de courtes animations de sensibilisation, comme celles sélectionnées dans la section des ressources ci-dessous. Ces animations sont courtes, divertissantes et souvent utiles pour faire passer des messages clés sur la préservation numérique à des publics non spécialisés et au grand public.

Ressources

Digital Preservation and Nuclear Disaster: An Animation
https://www.youtube.com/watch?v=pbBa6Oam7-w
Dessin animé divertissant sur l'importance des dépôts numériques fiables, des métadonnées et des rafraichissements de médias numériques (3 min 18 s).

Team Digital Preservation and the Aeroplane Disaster
https://www.youtube.com/watch?v=EKnsZZzuUr4
Dessin animé divertissant sur les effets de l'obsolescence et l'importance de la migration (3 minutes 37 secondes).

Team Digital Preservation and the Arctic Mountains Adventure
https://www.youtube.com/watch?v=PGFOZLecjTc
Dessin animé divertissant sur l'importance de la planification de la préservation (4 min 22 s).

Team Digital Preservation and the Deadly Cryptic Conundrum
https://www.youtube.com/watch?v=Yun9hkPPF9M
Dessin animé divertissant sur l'importance des informations de représentation (4 mn 9 s).

Études de cas

Increasing Participation in Internal RDM (research data management) Training Sessions
http://www.dcc.ac.uk/resources/developing-rdm-services/increasing-participation-training
Cette étude de cas examine les approches adoptées par deux projets MRD du Jisc pour assurer une bonne participation à leurs sessions de formation interne sur la gestion des données de recherche (RDM). 2013 (4 pages).

Defining and Formalizing a Procedure for Archiving
http://files.archivists.org/pubs/CampusCaseStudies/Case2Final.pdf
Nancy Deromedi, de l'Université du Michigan, décrit la formalisation d'un partenariat avec une unité administrative clé, qui n'avait pas été jusqu'à présent un partenaire réceptif sur le campus, et la sensibilisation aux considérations archivistiques alors que l'unité passait d'un système hybride de gestion numérique et papier à un processus uniquement numérique. Avril 2008 (8 pages).

Achats et recours à des tiers

Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Introduction

Cette section donne un aperçu des questions clés et des conseils pour la sélection et le recours à des services tiers dans le domaine de la préservation numérique. Les modalités d'acquisition d'une prestation varient souvent en fonction du secteur d'activité ou du pays. Les organisations doivent identifier et suivre leurs procédures d'achat légales et réglementaires pour s'assurer que l'achat des prestations suit les procédures correctes. Le non-respect de ces directives spécifiques pourrait entraîner d'importants problèmes, notamment une compensation pour d'autres prestataires potentiels désavantagés par des procédures d'achat incorrectes.

Trois tableaux sont fournis dans le cadre de ce manuel : Ressources humaines pour les tâches de gestion des achats ; Enjeux et potentiels avantages et inconvénients de l'utilisation de prestations fournies par des tiers dans les activités de préservation numérique et une liste des points à vérifier pour évaluer l'état de préparation du stockage pour la préservation numérique car l'achat tient souvent une place importante dans la mise en œuvre du stockage des archives (voir Services de stockage et Service *cloud*). La dernière section, Ressources, fournit des indications supplémentaires et une description sommaire d'autres orientations et études de cas.

Le coût sera évidemment un élément clé pour décider de sous-traiter ou non la préservation numérique, mais d'autres facteurs doivent être pris en compte et les avantages et inconvénients de chacun devront être envisagés par rapport à la mission globale de l'institution. Il s'agit notamment du contrat, de l'accord sur le niveau de service, de l'efficacité et de la qualité des services fournis, de l'intégration au sein des processus et de l'environnement de l'institution, des plans de reprise d'activité et de continuité des activités, de la possibilité de quitter le service si nécessaire et de la manière dont le service peut être contrôlé et mesuré. Les exigences légales en matière de confidentialité ou de protection des données peuvent par exemple influencer la pertinence de l'externalisation, compte tenu de la juridiction dont relève le fournisseur de services et de l'endroit où le service est physiquement situé.
L'externalisation de tâches ou de services spécifiques n'est en aucun cas un phénomène nouveau. Les entrepôts numériques sous-traitent certaines de leurs opérations depuis des décennies. C'est un domaine dans lequel les enseignements tirés de l'externalisation d'autres services peuvent être

utiles. Un enseignement majeur, directement applicable à l'environnement numérique, est l'importance cruciale de disposer de ressources en personnel et de connaissances suffisantes de la technologie pour pouvoir élaborer des spécifications efficaces.

Ressources humaines pour les tâches de gestion des achats

La mesure dans laquelle les avantages potentiels du recours à des prestations peuvent être maximisés et les inconvénients potentiels minimisés dépendra fortement de l'affectation de ressources en personnel aux activités suivantes :

• **Ressources humaines pour les tâches de gestion des achats**
Établir l'organisation et la gouvernance appropriées lors de la sélection des services de tiers : • Défendez le concept de préservation numérique. • Impliquez les parties prenantes internes appropriées dès le début de votre réflexion. • Développez une stratégie de communication non seulement pour votre équipe d'achat mais aussi pour un ensemble élargi de parties prenantes. • Tenez à jour un registre des risques pour la passation de marché. • Utilisez l'expertise au sein de votre organisation : par exemple, avez-vous un service « Achats »?
Établir des exigences claires et réalistes : • Alignez l'analyse de rentabilité sur les besoins identifiés au sein de l'organisation. Cela peut prendre un certain temps mais est essentiel pour obtenir un bon résultat. • Créez un environnement dans lequel vos parties prenantes peuvent contribuer au discours et sentir qu'elles ont eu leur mot à dire. • Apprenez de l'expérience des autres membres de la communauté de la préservation numérique dans son ensemble : cela pourrait inclure des visites de sites de référence et/ou le partage de documents et de points de vue. • Définissez les fonctions de préservation à inclure dans le marché, par exemple le versement, le stockage, la planification de la préservation, la préservation et les activités de gestion. Toutes les activités doivent-elles être sous-traitées à un prestataire ou font-elles simplement partie du cadre de la préservation numérique ? • Faites la distinction entre les fonctionnalités essentielles et les fonctionnalités à « valeur ajoutée » souhaitables : l'utilisation d'une méthodologie d'exigences particulières, par exemple le modèle du MoSCoW, apporte une discipline importante non seulement dans les premières étapes de la passation de marchés, mais aussi pour tout plan de projet à venir. • Ayez des exigences non ambiguës et mesurables que vous pouvez utiliser pour montrer clairement au prestataire s'il les respecte ou s'il n'est pas performant.
Clarifier les exigences légales : • Suivre les exigences et les processus institutionnels et réglementaires en matière de marchés. • La protection des données, la règlementation sur les données, les autres contenus sensibles et les droits d'auteur devront tous être pris en compte. • Cela vaut la peine de consacrer du temps à la négociation de vos contrats et accords avec des fournisseurs tiers. Les malentendus postérieurs prennent du temps. • Si des modifications du contrat ou l'accord proposé sont nécessaires, il est essentiel de les obtenir avant de signer tout contrat ayant force obligatoire en droit. Intégrez des possibilités de révision dans le contrat et comprenez quels sont les leviers dont vous

disposez. • Assurez-vous que vous n'êtes pas obligé d'attribuer un contrat pour conserver une certaine flexibilité jusqu'à cette attribution. • Assurez-vous que vous pouvez résilier le contrat de manière à perturber le moins possible les activités en cours si le prestataire ne respecte pas les exigences du contrat. Dans la mesure du possible, ne payez que les biens et services que vous avez reçus et non ceux que vous pourriez recevoir à l'avenir. • Comprenez clairement quels sont les services et produits offerts dans le cadre des coûts de base du contrat et quels sont ceux qui entraîneront des coûts supplémentaires. • Mettez toutes les exigences légales, y compris la juridiction/loi applicable, à la disposition des prestataires potentiels le plus tôt possible dans le processus de passation de marché, car cela peut avoir des conséquences importantes s'ils participent au processus.
Maintenir une bonne communication entre le prestataire et l'institution : • Accord de niveau de service pour identifier les rôles et les responsabilités de chaque partie. • Accès du personnel externe à l'infrastructure uniquement ou également à la maintenance et au développement ? • Construire des relations plus souples avec le fournisseur est également une compétence utile. • Existe-t-il, pour le système que vous avez choisi, une communauté d'utilisateurs active qui peut fournir des retours d'expérience et témoigne d'une bonne interaction avec le prestataire ?
Entreprendre des contrôles d'assurance qualité • Établissez la responsabilité pour des fonctions telles que le contrôle de l'intégrité. • Faites correspondre les contrôles d'assurance qualité avec les exigences mesurables que vous avez spécifiées dans le contrat et assurez-vous que le fournisseur réponde aux exigences ou change/corrige ses pratiques pour y répondre. • Assurez-vous de l'audit / respect des responsabilités légales.
Élaborer et suivre le contrat • Même si cela peut sembler prématuré, une stratégie de sortie doit être identifiée dès le départ. La fonction de préservation numérique durera plus longtemps que le fournisseur de services commerciaux et l'infrastructure technologique actuelle. • Comprenez bien vos droits concernant vos données. La récupération des données nécessaires à l'accès ou au transfert vers un autre fournisseur entraîne-t-elle des coûts ? • Soyez attentif au marché et aux modèles financiers utilisés par les fournisseurs. Vous devrez peut-être sortir des sentiers battus, car ces modèles ne correspondent pas forcément au modèle financier en vigueur dans votre organisation (les dépenses d'investissement par rapport aux dépenses de fonctionnement constituent un dilemme fréquent). • Ayez connaissance de tout changement de l'environnement technologique qui concernerait le fournisseur tiers. • Tenez-vous au courant du marché après avoir conclu votre achat. Vous devez savoir dans quelle mesure votre fournisseur est commercialement solide. Mettez périodiquement à jour les points de vérification de vos exigences.

Ces coûts devront être ajoutés aux coûts globaux du contrat lors du calcul du rapport coûts/bénéfices du recours à des services de tiers pour la préservation numérique, en gardant à

l'esprit que la plupart de ces coûts seront ou devraient être engagés même si la préservation n'était pas externalisée.

Enjeux et potentiels avantages et inconvénients de l'utilisation de prestations fournies par des tiers dans le cadre des activités de préservation numérique

Enjeu	**Avantage potentiel de l'utilisation d'une prestation fournie par un tiers.**	**Désavantage potentiel de l'utilisation d'une prestation fournie par un tiers.**
Un personnel, des compétences et une expérience limités	• Fournit des compétences et une expérience spécialisées qui peuvent ne pas être disponibles au sein de l'institution.	• Sans une certaine expérience pratique et une certaine expertise, il sera difficile d'élaborer et de contrôler des contrats efficaces. • Sans expérience pratique, il sera également difficile de comprendre et de communiquer efficacement les exigences de l'organisation (ou d'évaluer si elles sont techniquement réalisables ou non)
Coûts	• Évite de devoir développer des infrastructures coûteuses (particulièrement important pour les petites institutions). • S'il y a des économies d'échelle, l'externalisation peut être rentable.	• Il y a très peu de benchmarking établi. C'est un domaine encore trop récent. • Risque de faillite des entreprises. • Tant que le marché ne se développe pas, il peut y avoir une dépendance excessive à l'égard d'un seul prestataire.
Rapidité de déploiement	• Permet de prendre des mesures à court ou moyen terme, en attendant le développement des infrastructures.	• À moins qu'il n'y ait des stratégies de sortie adéquates, possibilité d'être bloqué dans un contrat d'externalisation plus longtemps que prévu.
Noyau de compétences	• Permet à l'institution de se concentrer sur d'autres services.	• Risque de ne pas développer ou de perdre les fondamentaux de compétences spécialisées. La capacité de prendre des décisions en connaissance de cause est indispensable.
Considérations relatives à l'accès	• Le suivi de l'utilisation peut être plus efficace (à condition que le prestataire ait démontré sa capacité à fournir des statistiques d'utilisation significatives). • Il peut y avoir des synergies et des économies de coûts en externalisant en même temps l'accès et la préservation.	• Des temps de réponse difficiles à contrôler, qui peuvent être inacceptables et/ou plus coûteux, en particulier pour les articles très utilisés • Il peut être difficile de prévoir les besoins futurs dans ce domaine.

Gestion des droits	• Évite ce qui est souvent une activité gourmande en ressources pour l'institution.	• Peut augmenter considérablement le coût du contrat et/ou compliquer les négociations avec les tiers détenteurs de droits.
Sécurité	• Le contrat peut garantir les dispositions de sécurité requises par l'institution.	• Manque de contrôle, en particulier pour des données sensibles.
Contrôle qualité	• Un contrat solide intégrera des exigences strictes en matière de contrôle de la qualité.	• Le risque de perte ou decorruption peut encore être jugé trop élevé pour des données hautement significatives et/ou sensibles.
Stockage	• Accès à des dispositifs de stockage éprouvés et gérés par des professionnels, avec reproduction facile du contenu et vérification de l'intégrité.	• Enjeux de confiance et considérations juridiques lors du stockage de données sensibles. • Il est difficile d'anticiper les coûts réels de certains services, par exemple le stockage et l'informatique dans le *cloud*, car l'organisation ne sait souvent pas exactement de quel service elle aura besoin. Cette incertitude peut être réduite avec l'expérience.

Liste de points de contrôle pour la sélection et la comparaison des prestataires de services

Les listes de points de contrôle et les normes peuvent être des points de départ précieux lorsque vous envisagez ou évaluez l'utilisation de services de tiers, car il s'agit de listes déjà établies que vous pouvez facilement adopter ou adapter à vos besoins. Les listes de points de contrôle vous aident notamment à identifier les éléments que vous pourriez autrement oublier de prendre en considération et à exprimer clairement les enjeux et les exigences.

Les listes de points de contrôle fonctionnent bien lorsqu'elles sont associées à un modèle de maturité. Par exemple, les niveaux de préservation de la NDSA permettent d'établir une liste de points de contrôle pour voir dans quelle mesure un prestataire de services répond à chaque niveau. Une organisation identifie le niveau de maturité dont elle a besoin maintenant et à l'avenir, puis recherche des prestataires de services dont les niveaux correspondent.

Les listes de points de contrôle et les normes relatives aux services d'entrepôts numériques sont des points de départ précieux, car vous pouvez choisir les parties de la liste de points de contrôle qui s'appliquent aux services spécifiques que vous recherchez. Des exemples de listes de points de contrôle et de normes pertinentes sont disponibles dans la section Ressources et sont également abordés plus en détail dans la section Audit et certification du manuel.

Ce manuel propose ci-dessous une liste de points de contrôle pour évaluer l'état de préparation du stockage pour la préservation numérique :

Liste de points de contrôle : questions à poser à votre prestataire de services de préservation

- ☐ Quel est le niveau de redondance du système de stockage ? À combien d'emplacements physiques distincts l'objet numérique est-il conservé ? Quelle est la distance géographique qui les sépare ?
- ☐ Différents types de technologies de stockage sont-ils utilisés pour atténuer/répartir les risques ? (par exemple, le stockage en ligne et hors ligne).
- ☐ Si un fichier a été corrompu ou modifié involontairement, comment et quand la détection intervient-elle ? Des journaux d'audit ou d'autres formes d'enregistrement sont-elles disponibles pour montrer que des contrôles d'intégrité des données ont été effectués et pour en indiquer le résultat ?
- ☐ Quelle est la stratégie de récupération en cas de catastrophe, par exemple si un système de stockage tombe en panne ou si une catastrophe naturelle se produit sur un site de stockage ? Comment les objets numériques sont-ils récupérés ? À quand remonte le dernier test de cette stratégie de reprise d'activité après sinistre?
- ☐ Quelle est la stratégie de migration du stockage pour faire face à l'obsolescence technique ? Que se passe-t-il lorsque le système est en fin de vie et que le contenu doit être migré vers un nouveau système ? Le contenu est-il encore accessible pendant ce processus ?
- ☐ Quelle est la stratégie de sortie lors de l'utilisation d'un type de stockage donné (par exemple, sur site, dans le *cloud*) ? Par exemple, que se passe-t-il si le vendeur du système de stockage fait faillite ?
- ☐ Quelles sont les mesures en place pour contenir les fichiers corrompus ou altérés, par exemple la mise en quarantaine des fichiers, pour éviter qu'ils ne soient répliqués ?
- ☐ Quelles sont les mesures de sécurité et d'audit mises en place pour empêcher l'accès et/ou la modification non désirés des objets numériques ?
- ☐ Qui est responsable du contrôle et de la gestion du système de stockage afin de garantir son bon fonctionnement ? Y a-t-il une continuité du personnel en cas de vacances, de maladie ou de départs ?
- ☐ Quels sont les contrats, garanties ou assurances accompagnant la solution ou le service de stockage qui engagent le vendeur ou le fournisseur à assurer la maintenance, la récupération ou le remplacement en cas de problème ?
- ☐ Quelle est l'approche ou la veille mise en place pour la surveillance des technologies de

	stockage et l'évaluation des risques afin que les migrations, les mises à jour, les mises à niveau ou la maintenance puissent être planifiées et exécutées en temps utile ?
☐	Les coûts et les risques sont-ils clairs afin qu'un compromis puisse être évalué et fait entre le nombre de copies, le type de stockage, la facilité d'accès et la sécurité des objets numériques ?
☐	Quelles sont les normes que le prestataire s'efforce de respecter (par exemple, OAIS, normes de sécurité de l'information) ? Vise-t-il à être reconnu comme un entrepôt numérique fiable ?
☐	Comment le prestataire peut-il démontrer qu'il fait ce qui a été convenu ?

Ressources

OAIS : *Open Archival Information Systems* : Modèle de référence pour un système d'information archivistique ouvert: Pratique recommandée (CCSDS 650.0-M-2 : livre magenta), Comité consultatif pour les systèmes de données spatiales 2012
https://public.ccsds.org/pubs/650x0m2.pdf (version anglaise)
https://public.ccsds.org/Pubs/650x0m2%28F%29.pdf (version française)
Fournit une terminologie et un modèle fonctionnel partagés utiles pour identifier les exigences relatives à l'acquisition de services de préservation numérique par des tiers.
Il a ensuite été publié sous la forme de la norme ISO 14721:2012 – Systèmes de transfert des informations et données spatiales – Système ouvert d'archivage d'information (OAIS) – Modèle de référence, 2e édition.

Core Trust Seal
https://www.coretrustseal.org/
Le sceau *Core Trust Seal* est la première étape du cadre mondial de certification des entrepôts numériques. Il comprend une liste de contrôle en 16 points et peut être utilisé pour l'auto-évaluation ou l'examen par les pairs.

ISO16363: 2012 Audit et certification des référentiels numériques de confiance (non disponible en français).
http://www.iso16363.org/
La norme ISO 16363 est un cadre d'audit pour la préservation numérique appuyé par de la documentation, qui comprend plus de 80 critères pouvant être utilisés pour l'auto-audit ou l'audit externe. Les critères utilisés dans la norme s'appliquent à l'ensemble de l'organisation et pas seulement au système technique dans lequel le contenu de la collection est stocké. La version pré-imprimée du *Magenta Book* de la norme (Comité consultatif pour les systèmes de données spatiales) est disponible gratuitement sur http://public.ccsds.org/publications/archive/652x0m1.pdf

DIN 31644 Information et documentation - Critères pour les archives numériques fiables (indisponible en français)

Explanatory notes on the nestor Seal for Trustworthy Digital Archives. Disponible à l'adresse suivante : http://files.dnb.de/nestor/materialien/nestor_mat_17_eng.pdf

Le processus d'auto-évaluation étendu pour les archives numériques est une liste de points de contrôle utile développée par nestor sur la base de la norme DIN 31644 Information et documentation - Critères pour les archives numériques fiables (44 pages).

The NDSA Levels of Digital Preservation: An Explanation and Uses

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

Les niveaux de préservation de la *National Digital Stewardship Alliance*(NDSA) sont largement utilisés dans le manuel et sont utiles pour réfléchir à de nombreux domaines de la préservation numérique. Il existe également des correspondances entre les niveaux de préservation de la NDSA et les profils des fournisseurs de stockage dans le *cloud* par AVPreserve (7 pages).

Voir la version française : Thomas Ledoux, Marion Humbert, Edouard Vasseur. *Utilisation des Niveaux de Préservation Numérique : aperçu de la V2.0*, traduction française. 2020.
https://hal-bnf.archives-ouvertes.fr/hal-02551807/document

Where to keep research data. DCC Checklist for Evaluating Data Repositories

http://www.dcc.ac.uk/sites/default/files/documents/publications/Where%20to%20keep%20research%20data.pdf

Une liste de points de contrôle utile du *Digital Curation Centre* sur les lieux où conserver les données de recherche en toute sécurité. Elle comprend un Accord de niveau de service avec des niveaux de maturité et concerne principalement les dépôts tiers externes qui offrent des prestations à la communauté britannique de la recherche (20 pages).

Le guide des Archives nationales du Royaume-Uni sur le stockage dans le *cloud*

http://www.nationalarchives.gov.uk/archives-sector/digital-collections.htm

Fournit des informations sur les marchés publics dans le contexte des services de *cloud computing* à des fins de préservation, y compris des études de cas de plusieurs institutions (voir ci-dessous). Il est particulièrement remarquable pour son examen des questions juridiques.

Journée d'étude de la DPC sur le sujet des achats en matière de préservation

http://www.dpconline.org/events/previous-events/1150-procuring-preservation-writing-and-understanding-requirements-in-digital-preservation

Pour un aperçu de certains des éléments d'exigences en matière de cadrage, voir les présentations individuelles énumérées. Les présentations sur l'analyse des exigences et les « Achats en matière de préservation : parcours d'obstacles et processus » sont particulièrement pertinents.

Études de cas

Groupe de travail sur la préservation numérique du *Archives&Records Council* du pays de Galles
http://www.nationalarchives.gov.uk/documents/Cloud-Storage-casestudy_Wales_2015.pdf
Cette étude de cas examine l'expérience d'un groupe de travail intersectoriel d'archives galloises qui a coopéré pour tester une série de systèmes et de déploiements de services dans le cadre d'un PoC pour l'archivage dans le *cloud*. Elle explique le contexte organisationnel, la nature variée de leurs exigences et approches en matière de préservation numérique, et leur expérience en matière de sélection, de déploiement et de tests de la préservation numérique dans le *cloud*. L'étude de cas a examiné le logiciel open source Archivematica avec Windows Azure de Microsoft ; Archivematica avec CloudSigma ; *Preservica Cloud Edition* et a commencé à tester Archivematica avec Arkivum 100. Janvier 2015 (10 pages).

Case Study: Tate Gallery[12]
http://www.nationalarchives.gov.uk/documents/Cloud-Storage-casestudy_Tate_Gallery_2015.pdf
Cette étude de cas traite de l'expérience de développement d'un système d'archivage électronique partagé pour les quatre sites physiques de la Tate, alimenté par un système de stockage commercial d'Arkivum. Elle explique le contexte organisationnel, la nature de leurs besoins et approches en matière de préservation numérique, et les raisons pour lesquelles ils ont choisi la solution sur site d'Arkivum, Arkivum/OnSite, de préférence à toute offre basée sur le *cloud*. Il se termine par les principaux enseignements tirés du projet par l'équipe et examine les plans de développement futur. Janvier 2015 (7 pages).

Case Study: Dorset History Centre
http://www.nationalarchives.gov.uk/documents/Cloud-Storage-case-study_Dorset_2015_%281%29.pdf
Cette étude de cas porte sur le *Dorset History Centre*, un service d'archives local du gouvernement [britannique (NdT)]. Elle explique le contexte organisationnel du service, la nature de ses exigences et approches en matière de préservation numérique, son projet pilote de deux ans utilisant *Preservica Cloud Edition* (un service de préservation numérique basé sur le *cloud*), l'infrastructure technique du système d'archivage électronique, ainsi que l'analyse de rentabilité et le financement du projet pilote. Il se termine par les principaux enseignements tirés du projet par l'équipe et une présentation des projets futurs. Janvier 2015 (9 pages).

Case Study: The Parliamentary Archives
http://www.nationalarchives.gov.uk/documents/Cloud-Storage-casestudy_Parliament_2015.pdf
Cette étude de cas couvre les Archives parlementaires et leur expérience de fonctionnement via le cadre G-Cloud. Pour plus de résilience/une stratégie de sortie, elles ont choisi deux fournisseurs de services dans le *cloud* avec différentes infrastructures de stockage sous-jacentes. Il s'agit d'un exemple de système d'archivage électronique utilisant un ensemble hybride de solutions de stockage (en partie dans le *cloud* public et en partie installé localement) pour la préservation numérique, car le SAE dispose d'un système de préservation installé localement (*Preservica*

[12] Ensemble composé de quatre musées d'art londoniens (NdT).

Enterprise Edition) qui est intégré au stockage dans le *cloud* et au stockage local et stocke les documents sensibles localement, et non dans le *cloud*. Janvier 2015 (6 pages).

Partnering with IT to Identify a Commercial Tool for Capturing Archival E-mail of University Executives at the University of Michigan
http://files.archivists.org/pubs/CampusCaseStudies/CASE-14-FINAL.pdf
Aprille Cooke McKay, de la *Bentley Historical Library* de l'Université du Michigan, examine les défis et les opportunités d'un partenariat avec les services informatiques pour lancer une consultation (*Request for Proposal* ou RFP) pour un logiciel commercial d'archivage de courrier électronique. 2013 (53 pages).

Digital Preservation Procurement Case Study
https://www.sheffield.ac.uk/library/digitalpreservation/casestudy
Un résumé du processus d'acquisition d'un système d'archivage électronique à l'Université de Sheffield (2 pages).

Audit et certification

Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Introduction

Les organisations s'intéressent de plus en plus à l'évaluation de leurs infrastructures de préservation numérique par rapport à un cadre d'évaluation, et l'audit, la certification et l'auto-évaluation sont des sujets brûlants dans le domaine de la préservation numérique. Il est intéressant de prendre un moment pour réfléchir à la différence entre un exercice d'auto-évaluation et un audit.

L'audit et la certification sont des processus formels généralement effectués et délivrés par des prestataires de services externes. Il s'agit souvent d'une expérience longue et exigeante qui démontre à un public externe qu'une norme particulière est respectée.

L'auto-évaluation est un précurseur, ou une alternative, à un audit complet et est généralement effectuée par le personnel de l'organisation. Les résultats sont généralement d'une grande valeur pour l'organisation évaluée (plutôt que pour un public externe). Les auto-évaluations peuvent être utiles pour identifier les pratiques insuffisamment développées et qui doivent être améliorées, en particulier si une organisation souhaite mener un audit complet et une certification à une date ultérieure.

L'assurance de disposer d'un entrepôt fiable est la garantie de nombreux avantages. Le concept d'entrepôt numérique fiable ou digne de confiance est désormais largement reconnu dans la communauté de la préservation numérique. La section suivante résume le travail qui a été effectué au cours des 10 à 15 dernières années pour en arriver là.

Contexte de l'élaboration des cadres d'audit et de certification

Les méthodes d'audit et de certification pour la mise en œuvre de la préservation numérique sont en cours d'élaboration depuis plus de dix ans, différentes organisations développant en parallèle différentes méthodologies. En Europe, ces méthodes sont maintenant regroupées dans le cadre européen pour l'audit et la certification des entrepôts numériques.

Le modèle de référence OAIS (ISO, 2012a) (voir Normes et bonnes pratiques) a influencé le développement des différentes méthodologies, qui a commencé avec la publication de *Trusted*

digital repositories : Attributes and responsibilities (RLG/OCLC, 2002). Ce document a été affiné sous la forme d'une ébauche de publication intitulée *An audit checklist for the certification of trusted digital repositories* (RLG-NARA, 2005) avant d'être finalisé sous la forme de *TRAC (Trustworthy Repositories Audit & Certification : Criteria and Checklist)* (CRL, 2007).

Des actions équivalentes se déroulaient également aux Pays-Bas et en Allemagne. Le processus d'auto-évaluation, « Sceau d'approbation des données » (*Data Seal of Approval*) développé par le DANS (*Data Archiving and Networked Services*), a été publié en 2008. Entre-temps, sur la base des recommandations d'un groupe de travail de nestor, le Comité allemand de normalisation (DIN) a adopté la norme DIN 31644 Information et documentation - Critères pour les archives numériques fiables.

Suite à la publication de la norme OAIS, et l'adoption ultérieure de l'OAIS comme norme ISO, en septembre 2011, le Comité consultatif pour les systèmes de données spatiales a publié des préconisations sur « l'audit et la certification des entrepôts numériques fiables » (*Audit and certification of Trustworthy Digital Repositories*), qui ont ensuite été adoptées et publiées sous la norme ISO 16363:2012 – *Audit et certification des dépôts numériques fiables* (ISO,2012b).

Options d'évaluation actuelles et cadre européen pour l'audit et la certification

La prolifération apparente des normes d'audit des entrepôts a été fréquemment citée comme un obstacle au lancement de projets. En conséquence, la Commission européenne a organisé une série de réunions pour convenir d'une approche à l'échelle européenne, et il existe maintenant un Protocole d'accord pour définir un cadre européen pour l'audit et la certification des entrepôts numériques. Ce protocole crée effectivement une approche à plusieurs niveaux de la certification, permettant une auto-évaluation et un examen par les pairs au niveau d'entrée, basés sur le sceau *CoreTrustSeal* (auparavant le sceau d'approbation des données), une auto-évaluation plus approfondie (basée sur la norme DIN 31644 ou ISO 16363), et un audit externe complet basé sur la norme ISO 16363.

1. *Data Seal of Approval* (Sceau d'approbation des données)

Le DSA, 2008 est un processus d'auto-évaluation pour les archives numériques, destiné spécifiquement aux services d'archives qui détiennent des données. Bien qu'il faille un certain temps pour demander le DSA, il est beaucoup moins onéreux que la norme ISO 16363, avec seulement seize lignes directrices sur lesquelles l'organisation est évaluée. Ces lignes directrices sont basées sur les cinq critères suivants :

- Les données peuvent être consultées sur Internet ;
- Les données sont accessibles (droits et licences clairs) ;
- Les données sont dans un format utilisable ;
- Les données sont fiables ;
- Les données sont identifiées de manière unique et persistante afin de pouvoir y faire référence.

Bien que le DSA soit en apparence un auto-audit, cet auto-audit est ensuite examiné par des pairs avant qu'un label ne soit attribué, ce qui ajoute un niveau d'autorité au processus. L'ouverture et la transparence sont encouragées et les institutions sont invitées à mettre leurs documents justificatifs (essentiellement de la documentation, des politiques et des procédures) gratuitement à disposition en ligne. Contrairement à un audit selon la norme ISO 16363, le pair examinateur

n'est pas tenu de se rendre dans l'institution pour constater que les politiques et les procédures fonctionnent dans la pratique. Le processus est donc très largement fondé sur la confiance.

À la suite du travail effectué avec le *World Data System* par le biais de la *Research Data Alliance*, des modifications des lignes directrices de DSA ont été proposées. Les dernières étapes de l'examen de ces modifications sont en cours. Les détails concernant la date et la manière dont la transition vers les nouvelles directives sera gérée seront publiés en temps utile, mais en attendant, le sceau actuel sera prolongé jusqu'en 2017.

2. DIN 31644 Information et documentation - Critères pour les archives numériques fiables

En Allemagne, le comité de normalisation DIN a adopté la norme DIN 31644 Information et documentation - Critères pour les archives numériques fiables, sur la base des recommandations d'un groupe de travail du réseau allemand de compétences pour la préservation numérique (nestor). La norme comprend des exigences pour un entrepôt numérique fiable structuré en trois sections :

Le cadre organisationnel exige que :

- L'entrepôt d'archives ait défini des objectifs pour la sélection des objets numériques et accepte la responsabilité de les préserver à long terme ;
- L'entrepôt dispose d'une communauté d'utilisateurs définie à laquelle seront fournis l'accès et la capacité de manier les objets numériques ;
- Les règles juridiques et contractuelles entre les créateurs de données et l'entrepôt d'archives numériques soient respectées ;
- Des structures organisationnelles suffisantes soient prévues en termes de personnel, de finances, de planification à long terme et de continuité du service ;
- Les processus et les responsabilités soient définis et documentés.

La gestion des objets exige que :

- L'intégrité et l'authenticité de l'objet numérique soient assurées ;
- Un plan stratégique pour les activités de préservation numérique soit mis en place ;
- Des paquets d'informations pour l'entrée, le stockage et la diffusion soient définis ;
- Une documentation adéquate soit fournie, y compris des identifiants pérennes et des métadonnées structurelles, techniques, de gestion et d'historique suffisantes ;
- L'objet numérique et les métadonnées connexes soient conservés ensemble en vue d'une préservation à long terme.

L'infrastructure et la sécurité exigent que :

- - L'infrastructure informatique puisse traiter l'objet numérique de manière adéquate et soit sécurisée.

La norme DIN 31644 est en allemand, mais une traduction anglaise est fournie par nestor sur son site web.

La procédure de certification étendue entreprise par nestor prend environ trois mois. Des conseils sur ce processus (Groupe de travail sur la certification de nestor, 2013) sont disponibles sur leur site web. Ce processus de certification ne doit pas être confondu avec un audit externe complet, qui nécessite une accréditation formelle selon la norme ISO 16363.

3. Audit ISO 16363 et certification d'entrepôts numériques de confiance

La norme ISO 16363 est un cadre d'audit basé sur des preuves qui utilise le terme « entrepôt » pour désigner l'organisation responsable de la préservation numérique plutôt que la seule infrastructure technique utilisée pour le stockage. Les critères utilisés dans la norme s'appliquent à l'ensemble de l'organisation et pas seulement au système technique dans lequel le contenu de la collection est stocké. Les mesures sont regroupées en trois domaines :

- **Infrastructure organisationnelle** : à savoir la gouvernance, la structure organisationnelle, la dotation en personnel, la responsabilité au sein des procédures, le cadre politique, la viabilité financière et les contrats, les licences et les responsabilités ;

- **Gestion des objets numériques** : à savoir la collecte et les entrées, la planification de la préservation, la création et la préservation des paquets d'informations archivistiques (AIP), et la gestion de l'information et de l'accès ;

- **Gestion des risques liés à l'infrastructure et à la sécurité** : à savoir l'infrastructure technique, la gestion des risques et la gestion des risques liés à la sécurité.

La terminologie utilisée dans la norme ISO 16363 est directement alignée sur celle de l'OAIS et la norme porte directement sur les paquets d'informations et les domaines fonctionnels de l'OAIS. Une compréhension de base de l'OAIS est donc utile pour ceux qui cherchent à comprendre la norme ISO 16363 et à fournir une évaluation par rapport à celle-ci.

Avec plus de 100 paramètres répartis sur les trois domaines, entreprendre un audit ou une évaluation de la norme ISO 16363 est un engagement important, similaire à de nombreuses autres normes ISO appliquées dans les organisations. Un nombre relativement restreint d'organisations ont utilisé la norme ISO 16363 depuis sa publication. Certaines ont demandé à être certifiées par des auditeurs externes, tandis que d'autres ont entrepris des auto-évaluations. Houghton (2015) reconnaît que même si une auto-évaluation n'est pas un audit, elle n'en reste pas moins un engagement important qui doit être adapté aux circonstances organisationnelles.

La norme ISO 16363 suit la pratique de l'ISO en matière de certification, qui suppose que les personnes effectuant l'audit sont elles-mêmes certifiées. Deux autres normes ISO fondent cette pratique :

- La norme ISO 16919 *Exigences pour les organismes d'audit et de certification des référentiels numériques potentiellement de confiance* (ISO, 2014) qui définit les exigences pour tout organisme qui certifie les auditeurs pour l'ISO 16363,
- La norme ISO 17021 *Exigences pour les organismes procédant à l'audit et à la certification de systèmes de management* (ISO, 2011 ; version française disponible) prévoit un mécanisme d'audit des organismes d'accréditation. Une version révisée a été publiée en 2015 : ISO/IEC 17021-1:2015 *Exigences pour les organismes procédant à l'audit et à la certification des systèmes de management — Partie 1: Exigences* (ISO, 2015; version française disponible).

Une agence du nom de PTAB (*Primary Trustworthy Digital Repository Authorisation Body*) propose des formations pour les auditeurs et ceux qui se préparent à l'audit. D'autres agences, dont le Centre des bibliothèques de recherche, proposent également des audits par rapport à ces normes.

4. Autres cadres et outils d'auto-évaluation

Les niveaux de préservation numérique de la NDSA (NDSA, 2013[13]) constituent une ressource utile pour les débutants. Ils sont particulièrement utiles pour les institutions qui débutent et peuvent être utilisés pour évaluer les premières étapes. Les niveaux de la NDSA sont largement utilisés dans le manuel (voir Pour commencer, Intégrité et empreintes, Sécurité de l'information et Stockage). Les cadres et les outils d'évaluation des risques peuvent également contribuer aux évaluations d'audit (voir Gestion des risques et du changement).

Quelle option d'audit ou d'évaluation dois-je choisir ?

Le Protocole d'accord de 2010 décrit ci-dessus, identifie effectivement une approche graduée de la certification. La quantité d'efforts requis pour chaque niveau augmente, mais le formalisme du résultat aussi. Le choix du cadre d'évaluation pour une organisation donnée doit donc prendre en considération au moins les éléments suivants :

Choisir un cadre d'évaluation	
Que voulez-vous obtenir de votre audit ?	Quel niveau de confiance essayez-vous d'atteindre ? Cherchez-vous à obtenir une certification d'une autorité externe, ou l'auto-évaluation est-elle suffisante ?
Quels sont les efforts ou les fonds disponibles pour réaliser l'évaluation ?	La norme ISO 16363 est une entreprise de grande envergure qui nécessite un effort important pour rassembler les preuves disponibles et réaliser l'audit ; le sceau *CoreTrustSeal* comporte beaucoup moins de paramètres et peut être réalisé dans un délai beaucoup plus court. La norme DIN 31644 propose deux options d'évaluation qui nécessitent des efforts variables.
Quel type de contenu votre institution possède-t-elle ?	À ce jour, le *CoreTrustSeal* a été principalement développé pour les entrepôts de données, tandis que le DIN et l'ISO 16363 sont tous deux neutres quant au contenu.
Quel cadre, le cas échéant, aura le plus de poids au sein de votre organisation ou auprès de vos parties prenantes externes ?	Y a-t-il une préférence nationale pour un cadre ou un cadre communément utilisé par des organisations similaires que vous devriez utiliser ?

Le choix du cadre d'évaluation ne doit pas être fait de manière isolée et peut souvent être directement influencé par la valeur qu'une évaluation peut avoir pour d'autres parties de l'organisation. Discuter des options avec les pairs et les responsables de l'organisation peut être une première étape utile pour s'assurer que la bonne option est choisie et que le soutien des autres secteurs de l'organisation est obtenu dès le départ.

Comment réaliser un audit ou une auto-évaluation ?

Après avoir choisi une méthodologie appropriée, une façon simple de procéder consiste à développer l'initiative en tant que projet et à produire un plan de projet. Les conseils sur la conduite de projet sont nombreux en ligne et il est utile de les consulter si votre organisation ne dispose pas d'un processus agréé pour la gestion de projet. Si elle en a un, vous devez vous

[13] Le lien proposé est celui de la V2 des niveaux, version française (avril 2020).

familiariser avec celui-ci et planifier votre projet en utilisant cette méthodologie (ou vous assurer l'assistance d'un gestionnaire de projet en interne). Votre plan doit comprendre au moins les sections suivantes :

- Périmètre : Quel est le contenu du périmètre de l'évaluation ?
- Calendrier : Quand l'évaluation aura-t-elle lieu et quand produira-t-elle des résultats ?
- Les parties prenantes : Qui réalisera l'évaluation ? Qui d'autre doit être interrogé ou consulté ?
- Gouvernance : Quel groupe assurera la gouvernance de l'évaluation et de ses résultats ?
- Communication : Comment le processus et les résultats seront-ils communiqués aux parties prenantes ?

Prochaines étapes : Comment les résultats seront-ils mis en œuvre ?

Si vous effectuez une évaluation ISO 16363, les conseils sur le site web de l'ISO 16363 Primary Trustworthy Digital Repository Authority Body vous seront utiles. La page de préparation de l'audit décrit les étapes à suivre pour réaliser un audit complet et ces étapes peuvent être adaptées pour une auto-évaluation. De même, le site web *CoreTrustSeal* dispose d'un outil d'auto-évaluation en ligne qui vous guidera tout au long d'une évaluation. Des versions PDF ou HTML des lignes directrices du manuel d'évaluation sont également disponibles.

Ressources

Rapport de l'APARSEN[14] sur l'examen par les pairs des entrepôts numériques
http://www.alliancepermanentaccess.org/wp-content/uploads/sites/7/downloads/2014/06/APARSEN-REP-D33_1B-01-1_1_incURN.pdf
Les leçons tirées à ce jour du processus de certification des entrepôts ont été utilement résumées par le projet APARSEN dans le présent rapport. Il suggère, bien que des progrès considérables aient été réalisés, que les procédures d'audit ne sont sans doute pas encore complètement ancrées et que certains problèmes subsistent, tant pour les auditeurs que pour les référentiels. (2012, 50 pages).

Outils de gestion de la préservation numérique : Principes
http://dpworkshop.org/workshops/management-tools/principles
Pour les organisations qui s'engagent à devenir des Entrepôts Numériques de Confiance, une étape d'acculturation pour développer un programme de préservation et de préservation numérique durable consiste à adapter et à adopter un ensemble de principes basés sur des normes. Ces principes fournissent un cadre pour votre programme et leur adoption est un point de départ positif (et, espérons-le, facile).

[14] Alliance Permanent Access to the Records of Science in Europe Network, réseau fondé en 2011 dans le but de créer à terme un centre virtuel d'excellence pour la préservation numérique (NdT).

Outils de gestion de la préservation numérique : Modèle de document

http://dpworkshop.org/workshops/management-tools/policy-framework

Chaque entrepôt numérique de confiance doit disposer d'un document de politique générale de haut niveau qui énonce explicitement la portée, le but, les objectifs, les principes de fonctionnement et le contexte du programme de préservation et de préservation numérique de l'organisation. L'équipe de l'atelier DPM a élaboré ce document modèle pour aider les organisations à atteindre cet objectif. Un document modèle identifie les sections recommandées d'un cadre de politique de préservation numérique avec des descriptions et des exemples pour chaque section.

Outils de gestion de la préservation numérique : auto-évaluation et audit d'examen par les pairs

http://dpworkshop.org/workshops/management-tools/self-assessment

Outil d'examen TRAC (*Trustworthy Repository Audit and Certification*) développé pour l'atelier DPM[15].

Open Archival Information System Reference Model: Introductory Guide (2nd Edition)

http://dx.doi.org/10.7207/twr14-02

Ce rapport de veille technologique de la DPC de 2014 fournit un guide succinct et accessible de la norme OAIS. La terminologie utilisée dans la norme ISO 16363 est directement alignée sur celle de l'OAIS. Le rapport sera utile pour une compréhension de base de l'OAIS utile pour comprendre l'ISO 16363 pour une évaluation par rapport à la norme.

Digital Preservation Capability Maturity Model (DPCMM)

http://www.securelyrooted.com/dpcmm/

Le DPCMM est un modèle de maturité étroitement aligné sur la norme OAIS. Il permet aux organisations d'évaluer leur état de préparation à la préservation numérique sur 5 niveaux pour 15 composantes.

CoreTrustSeal

https://www.coretrustseal.org/

Le sceau *Core Trust Seal* est la première étape du cadre mondial de certification des entrepôts. Il comprend une liste de contrôle en 16 points et peut être utilisé pour l'auto-évaluation ou l'examen par les pairs.

European Framework for Audit and Certification of Digital Repositories

http://www.trusteddigitalrepository.eu/

En 2010, le cadre européen pour l'audit et la certification des entrepôts numériques a été établi pour faire le lien entre la certification du sceau d'approbation des données (DSA), le groupe de travail sur l'audit et la certification des entrepôts de la CCSDS, et le groupe de travail des normes allemandes (DIN 31644) sur la certification pour les archives dignes de confiance. Il vise à maintenir un cadre intégré pour l'audit et la certification des entrepôts numériques et consistant en une séquence de trois niveaux, de plus en plus exigeants.

---

[15] *Digital Preservation Management Workshop*, groupe de réflexion sur la préservation numérique animé notamment par les bibliothèques du *Massachusetts Institute of Technology* (MIT Librairies) et l'Université Cornell. (NdT)

Études de cas

Preserving the H-Net Academic Electronic Mail Lists
http://files.archivists.org/pubs/CampusCaseStudies/Case11Final.pdf
Lisa M. Schmidt, de l'université d'État du Michigan, décrit l'évaluation de l'état de préservation actuel des listes de diffusion H-Net à l'aide de la théorie de la préservation numérique et de *L'Audit et la certification des dépôts de confiance : outil d'évaluation des critères et des points de contrôle (TRAC)*. Formuler des recommandations et superviser la mise en œuvre des améliorations pour faire de H-Net un dépôt numérique fiable. Garantir l'authenticité est la principale question de préservation. 2009 (15 pages).

ADS and the Data Seal of Approval - Case Study for the Digital Curation Centre
http://www.dcc.ac.uk/resources/case-studies/ads-dsa
Jenny Mitcham et Catherine Hardman, collègues du service des données archéologiques, décrivent l'expérience du SDA en matière de demande de label de qualité pour les données (DSA). Elles identifient des informations pratiques sur le processus de demande de DSA. Elles décrivent également les problèmes rencontrés par le SDA dans le cadre de ce processus et les avantages potentiels qu'elles envisagent de tirer de l'auto-certification du DSA. 2011.

Self-assessment of the Digital Repository at the State and University Library, Denmark – A Case Study
https://ipres-conference.org/ipres14/sites/default/files/upload/iPres-Proceedings-final.pdf
Dans ce document iPres 2014, les auteurs décrivent le processus et les avantages de réaliser un audit basé sur l'auto-évaluation et la norme ISO 16363 pour l'entrepôt numérique de la Bibliothèque nationale et universitaire du Danemark. (p.272-279 de 385)

TRAC Audit: Lessons
http://blog.dshr.org/2014/08/trac-audit-lessons.html
Voici le troisième article d'une série publiée sur le blog de David Rosenthal au sujet de l'audit TRAC des archives CLOCKSS[16] par le CRL[17]. Les articles précédents annonçaient la publication du rapport de certification et retraçaient le processus d'audit. Le présent billet examine les enseignements que CLOCKSS et d'autres peuvent tirer de leur expérience pendant l'audit.

Trustworthiness: Self-assessment of a Digital Repository against ISO 16363-2012
http://www.dlib.org/dlib/march15/houghton/03houghton.print.html
En 2013, la bibliothèque de l'Université Deakin a entrepris une auto-évaluation selon les critères de la norme ISO 16363. Cette expérience a abouti au présent rapport, qui fournit une évaluation de la norme ISO 16363, du processus d'évaluation et des conseils pour ceux qui envisagent de se lancer dans une entreprise similaire.

[16] *Controlled Lots of Copies Keep Stuff Safe*, système d'archivage électronique initié par l'Université de Stanford en 1999 (NdT).
[17] Centre for Research Libraries, réseau international de bibliothèques de recherche universitaires et indépendantes (NdT).

Managing an ISO 16363 Self-assessment: A How-To Guide
http://www.dcc.ac.uk/sites/default/files/documents/IDCC16/18_Managing_ISO16363.pdf
Une courte affiche présentée à la Conférence internationale sur la préservation numérique (IDCC) en 2016 par Maureen Pennock et Caylin Smith de la *British Library*.

Références

CRL, 2007. *Trustworthy Repositories Audit & Certification: Criteria and Checklist*. Disponible à l'adresse suivante : http://www.crl.edu/sites/default/files/d6/attachments/pages/trac_0.pdf

CCSDS, 2011.*Audit and certification of Trustworthy Digital Repositories*. Disponible à l'adresse suivante : http:// and certification of Trustworthy Digital Repositories.

DIN[18], 2012.*DIN 31644 Information and documentation – Criteria for Trusted Digital Repositories* (non disponible en français). Disponible à l'adresse suivante : http://www.nabd.din.de/cmd?level=tpl-art-detailansicht&committeeid=54738855&artid=147058907&languageid=de&bcrumblevel=3&subcommitteeid=112656173

Houghton, B., 2015. *Trustworthiness: Self-assessment of an institutional repository against ISO 16363-2012*. DLib Magazine, 21(3/4). Disponible à l'adresse suivante : http://www.dlib.org/dlib/march15/houghton/03houghton.html

ISO, 2014. *ISO 16919:2014– Systèmes de transfert des informations et données spatiales*–Exigences pour les organismes d'audit et de certification des référentiels numériques potentiellement de confiance (non disponible en français). Disponible à l'adresse suivante : http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=57950
ISO/IEC, 215. ISO/IEC 17021-1:2015 *Exigences pour les organismes procédant à l'audit et à la certification des systèmes de management – Partie 1: Exigences.* Disponible à l'adresse suivante :https://www.iso.org/fr/standard/61651.html

ISO, 2012a. *ISO 14721:2012 - Systèmes de transfert des informations et données spatiales – Système ouvert d'archivage d'information (OAIS) – Modèle de référence*, 2nd ed. Geneva: International Organization for Standardization. Disponible à l'adresse suivante : http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=57284

Version française : *Modèle de référence pour un système ouvert d'archivage d'information (OAIS), Pratique recommandée*, CCSDS 650.0-M-2 (F), Livre Magenta (juin 2012, version française d'octobre 2017). Disponible à l'adresse suivante :
https://public.ccsds.org/Pubs/650x0m2%28F%29.pdf

ISO, 2012b. *ISO 16363:2012 - Space data and information transfer systems – Audit and certification of trustworthy digital repositories*. Geneva: International Organization for Standardization. Disponible à l'adresse suivante :http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=56510

[18]Institut allemand de normalisation (NdT).

ISO/IEC 17021-1:2015 Exigences pour les organismes procédant à l'audit et à la certification des systèmes de management — Partie 1: Exigences. Disponible à l'adresse suivante : *https://www.iso.org/fr/standard/61651.html*

NDSA, 2013.*The NDSA Levels of Digital Preservation: An0Explanation and Uses, version 1 2013*. National Digital Stewardship Alliance. Disponible à l'adresse suivante : http://www.digitalpreservation.gov/ndsa/working_groups/documents/NDSA_Levels_Archiving_2013.pdf.

Version française : Thomas Ledoux, Jordan de la Houssaye, Stéphane Reecht, Bertrand Caron, Megan Phillips, et al..*NDSA Levels of Preservation* version 1, traduction française. 2019.https://hal-bnf.archives-ouvertes.fr/hal-02551807/document

nestor Certification Working Group, 2013. *Explanatory notes on the nestor Seal for Trustworthy Digital Archives*, nestor Materials 17, July 2013. Disponible à l'adresse suivante : http://files.dnb.de/nestor/materialien/nestor_mat_17_eng.pdf

RLG/OCLC[19] Working Group on Digital Archive Attributes, 2002. *Trusted digital repositories: Attributes and responsibilities*, Mountain View, California. Disponible à l'adresse suivante : http://www.oclc.org/content/dam/research/activities/trustedrep/repositories.pdf

RLG-NARA[20] Task Force on Digital Repository Certification, 2005.*An audit checklist for the certification of trusted digital repositories*, Mountain View. Disponible à l'adresse suivante : https://web.archive.org/web/20051126181100/http://www.rlg.org/en/pdfs/rlgnara-repositorieschecklist.pdf

[19]*Research Libraries Group*, réseau de bibliothèques américaines fondé en 1974 et fusionné en 2006 avec le *Online Computer Library Center*, organisation à but non lucratif dont l'objectif est d'offrir au public un meilleur accès aux informations (NdT).

[20] National Archives ana Records administration, Archives nationales américaines (NdT).

Respect de la législation

Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Introduction

Les informations fournies dans cette section sont uniquement destinées à donner des indications générales sur les questions juridiques découlant des divers aspects de l'archivage et de la préservation numériques et ne constituent pas un avis juridique. Il ne s'agit pas non plus de conseils sur les questions juridiques générales qui ont une incidence sur le fonctionnement des bibliothèques, des archives et d'autres entrepôts, car celles-ci sont traitées dans un certain nombre d'autres ouvrages de référence. Cette section est rédigée depuis une perspective britannique et la législation dans ce domaine varie d'un pays à l'autre. Bien qu'elle couvre principalement les questions juridiques britanniques et européennes, de nombreux sujets s'appliqueront également en termes généraux dans le droit d'autres pays.
Les informations fournies ne créent aucune relation conseiller-client. Si vous avez besoin de plus de détails concernant vos droits et obligations, ou de conseils juridiques sur les mesures à prendre, veuillez contacter un juriste ou un avocat.

Questions juridiques

« Ceux qui s'occupent de la préservation numérique doivent travailler dans le cadre de la loi telle qu'elle existe. Cela exige à la fois une bonne connaissance générale de la loi et un certain pragmatisme dans son application. Cette connaissance permet à l'archiviste d'éviter les pièges de la prudence excessive et d'une trop grande aversion pour le risque, et d'évaluer plus précisément les risques et les avantages de la préservation de nouvelles itérations de travaux numériques ». (Charlesworth, 2012, p.3)

Droits de propriété intellectuelle (DPI) et préservation

Les droits de propriété intellectuelle (DPI) sont une section du droit britannique qui comprend les brevets, les marques, les droits d'auteur et les droits associés - tels que le droit moral de l'auteur (voir Parties prenantes, conditions de contrat et de subvention) et les droits de représentation. La préservation des objets numériques nécessite souvent de planifier des séries d'opérations, ce qui crée des problèmes de DPI sans doute plus complexes et plus importants pour les objets numériques que pour les médias analogiques. Si elles ne sont pas traitées, ces questions peuvent entraver, voire empêcher les activités de préservation.

Qu'y a-t-il de différent avec le droit d'auteur et les objets numériques ?

Parmi l'éventail des DPI, le droit d'auteur revêt une importance particulière lorsqu'il s'agit d'envisager des actions de préservation numérique. La loi britannique sur le droit d'auteur a été élaborée en tenant compte des objets analogiques. Les objets analogiques traditionnels sont relativement stables, et des cadres juridiques et organisationnels pour la préservation sont bien établis. Le cadre juridique permettant d'entreprendre des travaux de préservation sur des objets numériques n'est pas aussi bien développé et les bonnes pratiques de préservation ne sont pas toujours reconnues ou autorisées par les dispositions existantes de la législation actuelle.

Le droit d'auteur fait une distinction entre la propriété de la manifestation physique d'une œuvre, comme un livre ou une œuvre d'art, et le droit distinct de la reproduire (droit de copie). Par nature, l'objet numérique ne correspond pas à cette distinction qui peut être source de confusion lorsqu'elle est appliquée sur le terrain. Dans le cas des objets numériques, les services de base d'un entrepôt, tels que la fourniture d'un accès aux utilisateurs et les activités de préservation de routine, impliquent souvent la création délibérée ou involontaire de copies. En l'absence d'autorisation de droits, de licences ou d'exceptions légales appropriées, ces copies peuvent constituer des violations du droit d'auteur. Le contenu numérique suppose donc un ensemble différent de considérations pour les entrepôts qui en détiennent. En outre, contrairement à l'objet physique, l'objet numérique nécessite la prise en compte de dépendances telles que le matériel et les logiciels, qui ont toutes leurs propres considérations distinctes en matière de propriété intellectuelle.

Une deuxième différence importante est la durée de vie commerciale et technologique relativement courte de l'objet numérique. La durée des DPI sur les objets numériques dépasse à la fois la « durée de préservation » commerciale et, dans presque tous les cas, la technologie dont ils dépendent. Il s'agit d'un triple problème, en termes d'obtention de licences pour la reproduction du contenu, de licences pour les logiciels d'accès au contenu et d'obtention des droits sur l'objet numérique « abandonné », sans compter l'urgence d'entreprendre ces actions.

Exceptions au droit d'auteur

Bien que la loi sur les droits d'auteur, les dessins et modèles et les brevets (1988) limite les actions de préservation possibles pour l'objet numérique, des exceptions pour les archives, les bibliothèques et les musées ont été introduites pour répondre à leurs exigences uniques de préservation. Du point de vue de la préservation, la disposition la plus importante (Office de la propriété intellectuelle, 2014) est le droit de produire le nombre de copies nécessaires à la préservation de l'objet numérique. Une autre exception importante est celle du terminal dédié, qui permet de mettre à disposition une copie numérique (c'est-à-dire l'une des copies créées dans le cadre de l'exception de préservation) accessible aux utilisateurs sur place. Ces dispositions ne s'étendent qu'aux éléments détenus en permanence dans la collection. Les exceptions permettent aux institutions couvertes par les exemptions de conserver des copies de documents sous différents formats de fichiers et d'adhérer ainsi à ce qui est considéré comme une bonne pratique de préservation tout en restant dans le cadre de la loi. Il convient de noter que les dispositions relatives aux exceptions au droit d'auteur n'annulent pas le droit moral de l'auteur qui doit toujours être pris en compte lors de la réalisation de travaux de préservation.

L'exemption de copie à des fins de préservation ne tient toutefois pas compte de la nature dépendante de l'objet numérique, et les dépendances à des logiciels tiers peuvent toujours constituer un obstacle aux actions de préservation. C'est un problème particulièrement observé dans les stratégies de préservation qui reposent largement sur la préservation de l'environnement

technique plus large des objets numériques en question. Par exemple, l'émulation en tant que stratégie de préservation nécessite l'utilisation de systèmes d'exploitation et de logiciels originaux externes à la collection permanente de l'entrepôt (voir Mesures de préservation). Il est important de tenir compte des coûts et du temps supplémentaires qui sont nécessaires pour entretenir des relations avec les tiers détenteurs de droits qui découlent de dépendances non couvertes par les exceptions.

Œuvres orphelines

Pour les institutions qui cherchent à publier des substituts numériques de contenu analogique, le programme d'octroi de licences pour les œuvres orphelines géré par l'Office de la propriété intellectuelle du Royaume-Uni, ainsi que l'exception européenne pour les œuvres orphelines sont susceptibles d'avoir un impact sur le travail de numérisation et la planification. Le programme de licence permet la numérisation commerciale et (dans le cas des institutions patrimoniales) non commerciale de tout type de contenus dont il n'a pas été possible de retrouver les détenteurs de droits à la suite d'une « recherche diligente ». La licence est un régime payant limité à une période de sept ans et destiné à être utilisé exclusivement au Royaume-Uni. Les dépositaires doivent planifier et budgétiser le renouvellement de ces licences. L'exception européenne pour les œuvres orphelines, en revanche, est limitée aux œuvres textuelles et audiovisuelles uniquement (et aux œuvres artistiques tant qu'elles sont intégrées dans les premières), ainsi qu'aux musées, bibliothèques, archives, établissements d'enseignement et radiodiffuseurs publics. Dans ce cas, l'avantage est que les recherches diligentes sont auto-certifiées et que la copie de préservation de l'œuvre, créée dans le cadre de l'exception de préservation, peut être mise en ligne pour une utilisation non commerciale, par exemple, ce qui facilite grandement les activités de numérisation.

Accès et sécurité

Une partie de la complexité supplémentaire des questions de droit d'auteur est également liée au fait que les objets numériques sont faciles à copier et à redistribuer. Les titulaires de droits sont donc particulièrement soucieux de contrôler l'accès et les éventuelles violations du droit d'auteur. Les technologies de gestion des droits numériques (DRM) développées pour répondre à ces préoccupations et fournir des mesures en matière de droit d'auteur, telles que les logiciels de protection contre la copie des fichiers et les erreurs physiques intentionnelles sur les CD/DVD, peuvent inhiber ou empêcher les actions nécessaires à la préservation. Les technologies DRM sont également elles-mêmes sujettes à obsolescence. Ces préoccupations concernant l'accès et la violation des droits doivent être comprises par les organisations qui conservent les objets numériques lorsqu'elles négocient des accords de dépôt avec les titulaires de droits, et prises en compte par les deux parties lors de la négociation des droits et des procédures de préservation. La mise en place de procédures de dépôt claires peut atténuer les problèmes d'accès futurs (voir Négociation des droits).

Archives web et dépôt légal

Le statut juridique des archives web et des processus de dépôt légal électronique varie d'un pays à l'autre : certains gouvernements ont adopté une législation sur le dépôt légal mais limitent l'accès aux seules salles de lecture des bibliothèques. Dans d'autres pays, il n'existe pas de législation sur le dépôt légal et les collections sont soit constituées uniquement sur la base d'une sélection et d'autorisations, soit conservées dans des « *dark archives* » inaccessibles au public. Au Royaume-Uni, les bibliothèques de dépôt légal ont le droit de rassembler et de donner accès à des copies de tous les sites web publiés dans le domaine britannique. Toutefois, l'accès à la collection est limité aux salles de lecture des bibliothèques (voir Milligan, 2015). Parallèlement, les archives du web

gérées par les Archives nationales du Royaume-Uni fonctionnent avec un champ d'application plus restreint concernant les publications gouvernementales et des pouvoirs statutaires plus clairs découlant de la législation sur les documents publics (voir Autres exigences légales).
Internet Archive, basé aux États-Unis, probablement les plus grandes et les plus utilisées des archives du web, n'ont aucune autorisation législative explicite pour récolter des sites web ou de les publier. Elles fonctionnent selon le principe « silence vaut accord », en supprimant de leurs collections tout site web dont le propriétaire demande le retrait. En revanche, la Bibliothèque du Congrès fonctionne sur la base d'une autorisation, ce qui signifie qu'elle doit obtenir l'accord explicite des détenteurs de droits d'auteur avant de récolter ou de publier leur contenu.

Autres exigences légales

D'autres exigences légales peuvent également s'appliquer et influencer la préservation des ressources numériques.

Les exigences de la législation sur les documents publics et les attentes connexes de la loi sur la liberté de l'information s'appliquent aux documents gouvernementaux, y compris ceux sous forme numérique. Les délais de préservation légaux et réglementaires s'appliquent à de nombreux documents numériques (par exemple, à des fins comptables et fiscales). Bien que ces périodes soient souvent limitées, il est à noter que les exigences de préservation des documents numériques dans certains secteurs (par exemple, l'industrie pharmaceutique, l'aide sociale et les dossiers médicaux) sont de plus en plus longues. Dans de tels cas, des stratégies de préservation à long terme s'appliqueront, car l'évolution technologique aura presque certainement une incidence sur l'accès à ces documents.

Les informations peuvent être soumises aux lois sur la protection des données et à la législation pertinente sur la protection de la vie privée protégeant les informations détenues sur les personnes. Au Royaume-Uni, le bureau du commissaire à l'information supervise le respect des questions de protection des données et de la vie privée.

Les informations peuvent également faire l'objet d'accords de confidentialité. Les préoccupations en matière de vie privée et de confidentialité peuvent avoir une incidence sur la manière dont les objets numériques peuvent être gérés au sein du dépôt ou par des tiers, et rendus accessibles pour être utilisés. La législation sur la protection des données a également un impact sur la circulation des données en dehors de l'Europe - une considération importante pour les organisations qui investissent dans un espace serveur à l'étranger.

Les décisions de l'Union Européenne sur le droit d'un individu à faire retirer ses informations personnelles des moteurs de recherche sur Internet dans certaines circonstances ont un impact significatif sur les pratiques des organisations travaillant avec des contenus numériques provenant du web (Koops, 2011). L'obligation d'éviter de porter préjudice aux individus lors de la sauvegarde de leurs données sur de longues périodes se reflète dans le principe du droit à l'oubli, par la mise en œuvre de l'article 12 de la directive 95/46/CE dans la jurisprudence de plusieurs pays européens.

Parties prenantes, conditions de contrat et de subvention

Certains objets numériques sont le résultat d'un substantiel investissement financier public (par exemple, les conseils de recherche) et/ou d'éditeurs, et d'un investissement intellectuel de la part de chercheurs et d'auteurs individuels. Chacune de ces parties prenantes peut avoir un intérêt

dans la préservation ; l'organisation qui les préserve devra en obtenir les autorisations afin de sauvegarder et d'optimiser l'investissement financier ou la valeur intellectuelle et culturelle de l'œuvre pour les générations futures. Ces intérêts peuvent se manifester par des conditions de contrat, de licence et de subvention ou par des dispositions légales telles que le « droit moral » des auteurs.

Investissement du dépositaire dans les contenus déposés

Les détenteurs du contenu sur plusieurs décennies devront presque certainement investir des ressources pour générer une documentation et des métadonnées révisées et générer de nouvelles formes de contenu si l'accès doit être maintenu. Il faut prévoir d'autres questions de DPI dans ce nouvel investissement et envisager la réutilisation future de ces contenus. Lorsqu'un déposant ou un donneur de licence conserve le droit de retirer des documents des archives et qu'un investissement important dans ces documents pourrait être anticipé au fil du temps par l'institution détentrice, des frais de retrait pour compenser tout investissement peuvent être intégrés dans les accords de dépôt (Voir Négociation des droits).

Gestion des droits

Comme indiqué dans la section Questions juridiques, il est important que les questions de licence, de droit d'auteur et tout autre droit de propriété intellectuelle sur les ressources numériques à préserver soient clairement identifiées et que les conditions d'accès soient convenues avec le déposant et/ou les titulaires de droits. Si la propriété juridique de ces droits n'est pas claire ou est excessivement fragmentée, il peut être difficile de préserver les contenus et de permettre aux utilisateurs d'y accéder. La gestion des droits doit donc être abordée dans le cadre du développement des collections et des procédures d'accès et être intégrée aux stratégies institutionnelles de préservation. Le degré de contrôle ou la portée de la négociation que les institutions auront sur les droits varieront, mais dans la plupart des cas, les stratégies institutionnelles dans ce domaine aideront à orienter les procédures opérationnelles. Elle sera également un élément crucial de toute métadonnée de préservation (voir Métadonnées et documentation) et de toute modalité d'accès (voir Accès).

Négocier les droits

À mesure que le volume des objets numériques augmente et que la complexité des droits et le nombre de titulaires de droits sur les objets numériques ne cessent de croître, les négociations ad hoc entre les organismes de préservation et les déposants et entre les titulaires de droits eux-mêmes deviennent plus onéreuses et moins efficaces. Cela est particulièrement problématique pour les organisations ou les activités britanniques qui ne sont pas couvertes par les nouvelles exceptions au droit d'auteur.

L'élaboration de lettres types pour la cession, par les personnes employées, de leurs droits, de modèles d'accords de dépôt et de modèles de licences et de clauses couvrant les activités liées à la préservation contribue à rationaliser et à simplifier ces négociations. Les institutions doivent demander l'aide d'un juriste pour rédiger ces modèles et fournir des conseils au personnel sur la mise en œuvre ou les variations autorisées dans les négociations avec les détenteurs de droits.
Un certain nombre d'institutions ont développé des modèles qui peuvent être adoptés ou adaptés à des institutions et des exigences spécifiques. Les procédures décrites ci-dessous sont une synthèse des bonnes pratiques actuelles.

Procédures recommandées

- Ayez recours à une expertise juridique pour définir votre politique de gestion des droits et pour élaborer des documents.
- Élaborez des lettres types de cession de droits, des accords types de dépôt, des licences types et des clauses pour les activités de préservation.
- Si vous accordez des licences pour du contenu provenant de tiers, assurez-vous que ceux-ci prennent en compte l'accès futur au contenu concerné par la licence et qu'ils disposent de procédures solides pour le maintenir.
- Préparez des arguments et des explications raisonnées pour vos activités de préservation qui conviennent aux parties prenantes externes telles que les détenteurs de droits. Ceux-ci devront être convaincus de la nécessité de la préservation et persuadés que leurs intérêts seront préservés. N'oubliez pas que leur sensibilisation aux questions de préservation peut être faible.
- Conservez les dossiers détaillés des négociations sur les droits. Établissez un calendrier indiquant clairement la liste des contenus déposés et couverts par la licence. Cela permettra de s'assurer que tout ce que l'on pense avoir été envoyé par le déposant a été reçu et pourra constituer la base d'un accusé de réception.
- Conservez les informations sur les droits et les autorisations pour tous vos objets numériques. Traitez les licences, les annexes et la correspondance relative aux droits comme des documents institutionnels clés à conserver dans des environnements à l'épreuve du feu et sécurisés.

Résumé des questions relatives aux licences et aux accords de dépôt

Vous trouverez ci-dessous une brève liste de points de contrôle et un résumé des questions juridiques qui peuvent être nécessaires à prendre en compte en ce qui concerne les licences de préservation ou les accords de dépôt d'objets numériques. Les exigences diffèrent selon les institutions, les secteurs et les pays, et la liste doit être adaptée aux besoins individuels. Cette liste ne constitue pas un conseil à valeur légale. Consultez un juriste pour traiter votre situation particulière.

DPI et préservation numérique

Une clause devrait être rédigée pour couvrir les points suivants :

- - Autorisations nécessaires pour le contenu.
- - Autorisations nécessaires pour les logiciels associés.
- - Autorisations nécessaires pour la copie à des fins de préservation. (Cette section s'applique au contenu ou aux organisations qui ne sont pas couverts par les exceptions au droit d'auteur)
- - Autorisations nécessaires pour la migration future du contenu vers de nouveaux formats à des fins de préservation. (Cette section s'applique aux contenus ou organisations non couverts par les exceptions au droit d'auteur)
- - Autorisations nécessaires pour l'émulation à des fins de préservation.
- - Autorisations relatives aux mécanismes de protection des droits d'auteur.

Accès

- - Autorisations et conditions relatives à l'accès à l'objet.

Questions statutaires et contractuelles

- Autorisations légales et obligations de dépôt légal en ce qui concerne les objets numériques.
- Subventions et obligations contractuelles concernant les objets numériques.

- Conditions, droits et intérêts appropriés des auteurs, éditeurs et autres financeurs.
- Informations confidentielles et protection de la confidentialité des personnes et des institutions.
- Protection de l'intégrité et de la réputation des créateurs de données ou d'autres parties prenantes.

Investissement par l'institution de préservation

- DPI sur toute valeur ajoutée par l'institution de préservation.
- Clauses de retrait (et frais associés).

Ressources

La législation britannique est régulièrement modifiée. Pour vous assurer que vous accédez aux dernières mises à jour, veuillez vous référer directement à : http://www.legislation.gov.uk/

Exceptions to Copyright: Libraries, archives and museums
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/375956/Libraries_Archives_and_Museums.pdf
Ce guide publié par l'Office de la propriété intellectuelle en 2014 présente les exceptions applicables aux bibliothèques, archives et musées. Il s'adresse à tous ceux qui travaillent dans ou avec des bibliothèques, des archives ou des musées au Royaume-Uni, ou à des détenteurs de droits d'auteur dont le contenu est détenu par ces institutions. Il couvre deux modifications importantes de la législation britannique qui touchent les bibliothèques, les archives et les musées. La première concerne la réalisation de copies d'œuvres afin de les préserver pour les générations futures. La seconde permet une plus grande liberté de copier des œuvres pour ceux qui effectuent des recherches non commerciales et des études privées.

Intellectual Property Rights for Digital Preservation
http://dx.doi.org/10.7207/twr12-02
Ce rapport de veille technologique de la DPC a été publié par Andrew Charlesworth en 2012. Le document ne couvre pas les législations récentes (telles que le règlement de 2013 sur les bibliothèques de dépôt légal [œuvres non imprimées] et les exceptions au droit d'auteur de 2014 pour les bibliothèques, les archives et les musées), mais constitue par ailleurs un ouvrage d'introduction pertinent. Le rapport s'adresse principalement aux déposants, aux archivistes et aux chercheurs/réutilisateurs d'œuvres numériques. Le droit de la propriété intellectuelle, représenté principalement par le droit d'auteur et ses droits connexes, a été de loin la branche du droit la plus prégnante, et souvent insoluble, en matière de préservation numérique. Il est essentiel que ceux qui s'engagent dans la préservation numérique soient capables d'identifier et de mettre en œuvre des stratégies pratiques et pragmatiques pour gérer les risques juridiques liés aux droits de propriété intellectuelle dans la poursuite des objectifs de préservation et d'accès (54 pages).

Aligning National Approaches to Digital Preservation
https://educopia.org/wp-content/uploads/2018/06/Aligning_National_Approaches_to_Digital_Preservation.pdf

Ces actes de 2012 comprennent deux documents sur des questions juridiques : *Legal Alignment* par Adrienne Muir, Dwayne Buttler et Wilma Mossink (p. 43-74) ; et *Legal Deposit and Web Archiving* par Adrienne Muir (p. 75-88). Le premier document se concentre sur les questions clés du dépôt légal, des exceptions au droit d'auteur pour la préservation et l'accès, et du travail et de la gestion des droits avec plusieurs partenaires et par-delà les frontières ; le second document aborde les défis de l'adaptation du dépôt légal, un mécanisme conçu pour l'édition imprimée, à l'environnement numérique. Les approches nationales des éléments importants de cadrage du dépôt légal et les questions juridiques découlant des approches non réglementaires de la collecte des publications numériques pour la préservation à long terme sont identifiées (342 pages).

Annexe du Guide sur le stockage dans le *cloud*, Tableau 3 - Questions juridiques
http://www.nationalarchives.gov.uk/documents/CloudStorage-Guidance_March-2015.pdf
Le tableau 3 fourni dans la section 7 en annexe du Guide sur le stockage dans le *cloud* des Archives nationales du Royaume-Uni publié en 2015, énumère les points juridiques de manière plus détaillée pour chacune des trois catégories clés :

- Les exigences légales en termes de gestion, de préservation et d'accès imposées aux Archives et à leurs organisations mères, par leurs donateurs et financeurs via des contrats et accords ou par la législation (par exemple, accessibilité, disponibilité, sécurité des informations, préservation, audit et conformité, loi sur les documents publics, etc.)
- les obligations légales relatives aux droits des tiers sur les données à stocker (par exemple, droits d'auteur, protection des données) ; et
- Les éléments juridiques de la relation entre un service d'archives et un ou plusieurs fournisseurs de services dans le *cloud* (par exemple, les conditions des contrats de service et les accords de niveau de service).

CREATe Working Paper No.3[21] *- Archives and Copyright: Risks and Reform n°3*
http://www.create.ac.uk/wp-content/uploads/2013/04/CREATe-Working-Paper-No-3-v1-1.pdf
Les pages 6 à 18 de cet article de R. Deazley et V. Stobo couvrent le droit d'auteur et le secteur des archives au Royaume-Uni (58 pages, 2013).

Rapport E-ARK sur les questions juridiques : la préservation culturelle en Europe dans un paysage législatif en mutation
http://www.eark-project.com/resources/project-deliverables/33-d22-legal-issues-report-european-cultural-preservation-in-a-changing-legislative-landscape
Ce rapport de 2017 de David Anderson, Université de Brighton, et du projet E-ARK[22] examine la législation en Europe et son impact sur la préservation numérique. La partie principale de ce rapport couvre les exigences de la directive 95/46/CE, qui ont été mises en œuvre par les États membres dans une variété d'instruments législatifs depuis l'adoption de la directive en 1995. Ces exigences sont mises en parallèle avec les propositions de règlement général sur la protection des données (RGPD) actuellement en discussion entre la Commission, le Conseil des ministres et le Parlement européen. La forme finale du texte n'ayant pas encore été arrêtée au moment de la rédaction du présent rapport, certaines de ses conclusions sont nécessairement provisoires... (130 pages).

[21] *Copyright and Creative Economy Centre,* organisme britannique basé à l'université de Glasgow.
[22] Projet européen développé entre 2014 et 2018 autour des questions de préservation numérique.

Legalities Lifecycle Management
https://www.timbusproject.net/portal/domain-tools/334-lehalities-lifecycle-management-tool/
Un outil développé par le projet TIMBUS qui s'intéresse à la préservation numérique des processus commerciaux. Les domaines couverts sont les DPI, les contrats informatiques, la protection des données et d'autres exigences légales.

Numérisation de masse du patrimoine culturel : les obstacles au droit d'auteur peuvent-ils être surmontés ?
http://livestream.com/unc-sils/iPres-Pamela-Samuelson/videos
Présentation principale d'iPRES 2015 par Pamela Samuelson, professeur de droit et d'information à l'Université de Californie, Berkeley. Mme Samuelson a publié de nombreux ouvrages sur les DPI et le cyberdroit. Dans cette présentation, elle examine le rôle de l'« utilisation équitable » dans l'approche du défi que représente le droit d'auteur. Samuelson s'exprime dans une perspective juridique américaine, mais de nombreuses considérations s'appliquent également au contexte britannique (2015, 56 minutes).

Références

Charlesworth, A.J., 2012. *Intellectual Property Rights for Digital Preservation*, *DPC Technology Watch Report 12-02*. Disponible à l'adresse suivante : http://dx.doi.org/10.7207/twr12-02

Intellectual Property Office, 2014. *Exceptions to Copyright: Libraries, Archives and Museums*. Disponible à l'adresse suivante :
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/375956/Libraries_Archives_and_Museums.pdf

Koops, B., 2011.*Forgetting Footprints, Shunning Shadows. A Critical Analysis of the "Right To Be Forgotten" In Big Data Practice*. SCRIPTed, 8:3, 229-256. Disponible à l'adresse suivante :
http://script-ed.org/wp-content/uploads/2011/12/koops.pdf

Milligan, I., 2015. *Web Archive Legal Deposit: A Double-Edged Sword*. Disponible à l'adresse suivante :
http://ianmilligan.ca/2015/07/14/web-archive-legal-deposit-a-double-edged-sword/

Gestion des risques et du changement

Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Introduction

La préservation numérique n'est pas seulement une question de risques. Elle crée également des opportunités et, en protégeant les objets numériques, elle signifie que l'on peut en tirer une valeur nouvelle ou étendue. Il peut être facile de se laisser submerger par les risques, il est donc utile d'être explicite dès le début du processus sur les opportunités qui sont protégées ou créées. De nombreux facteurs mettent vos ressources numériques en danger, notamment les changements apportés à votre organisation ou à votre technologie. S'ils ne sont pas gérés, ces risques auront un impact significatif sur votre capacité à mener des activités de préservation numérique, sur le périmètre de votre secteur d'activités ou sur votre capacité à vous conformer à la législation.

Pour gérer la préservation numérique, vous devez comprendre les problèmes et les risques spécifiques de votre organisation. Vous pouvez le faire en procédant à une évaluation des risques et des opportunités. Cette évaluation mettra en évidence les risques spécifiques pour la continuité de vos ressources numériques, ainsi que les opportunités créées en atténuant ces risques.

La gestion des risques

L'expérience montre que les risques auxquels sont exposées les ressources numériques sont subtils et variés. Ils comprennent, sans s'y limiter, les éléments suivants :

- Fusion, fermeture ou transfert de fonctions entre organisations.
- Changements d'orientation stratégique ou de financement et fonctions assurées par une organisation.
- Changements majeurs de dirigeants ou d'experts individuels.
- Externalisation sans prise en compte des besoins futurs de préservation.
- Obsolescence des formats de fichiers, ce qui signifie que le
- Défaillance des moyens de recherche entraînant des difficultés de récupération des données.
- Perte de droits d'auteur ou d'autres informations juridiques entraînant une incertitude sur les droits et obligations.
- Perte d'informations sur la provenance ou la fiabilité d'un document entraînant une perte d'authenticité.
- Défaillance du contrôle de version rendant difficile l'identification des instances faisant autorité pour un document.
- Erreur humaine entraînant une suppression accidentelle.
- Degré d'utilisation. Une archive peu consultée est plus à risque qu'une archive fortement utilisée. Si

traitement des données est coûteux ou impossible.
- Obsolescence des supports, qui rend la récupération des données coûteuse ou impossible.
- Dégradation des supports, qui signifie que les données sont endommagées ou modifiées.
- Perte d'informations contextuelles entraînant une perte de sens.

l'objet numérique est consulté peu fréquemment, l'impact de la défaillance est moins immédiatement visible.
- Catastrophes naturelles affectant les bâtiments ou les infrastructures.

La perte de données est susceptible d'avoir diverses conséquences concrètes, selon le contexte. Dans le contexte d'une affaire judiciaire, par exemple, l'authenticité d'un document peut devenir un problème juridique important ; pour des données de recherche très structurées en revanche, la chaîne de confiance peut être moins importante que l'accès à un contexte explicatif permettant la reproductibilité d'une expérience. Dans de nombreux contextes, il peut être techniquement possible de récupérer des collections numériques, mais lorsqu'une organisation n'a tout simplement pas les moyens ou les compétences nécessaires pour restaurer un ensemble de données, il peut en résulter une obsolescence pratique et une perte de données. Cela risque de devenir de plus en plus une réalité à mesure que le nombre et la complexité des collections numériques augmentent.
Les risques liés au contenu numérique sont généralement importants en raison de leurs conséquences concrètes. Là encore, cela dépend du contexte, mais peuvent se produire :

- Perte de réputation
- Ressources insuffisantes pour une tâche critique
- Incapacité à soutenir les utilisateurs dans leurs activités
- Défaillance pour remplir une fonction juridique ou réglementaire
- Incapacité d'exploiter et de réutiliser les données
- Perte d'identité et de mémoire dans l'institution
- Coût de la reconstitution et de la récupération

Les risques sont généralement classés par ordre de priorité en calculant un « score de risque » basé sur la probabilité, l'impact et l'imminence : un risque imminent avec une forte probabilité et un impact négatif important nécessite une action rapide. Selon la nature du risque, il peut s'agir de prendre des mesures pour réduire la probabilité qu'un risque apparaisse, de réduire l'impact si un risque se produit, ou de gagner du temps pour mettre en œuvre des mesures d'atténuation.

L'évaluation des risques est un processus continu qui peut être développé et étendu au fil du temps. Elle peut contribuer à réunir les différentes parties prenantes et, comme la gestion des risques est comprise par les instances dirigeantes, elle peut également contribuer à justifier les investissements. Même une évaluation élémentaire des risques mettra en évidence les priorités de quiconque se lance dans la préservation numérique.

Enfin, il convient de noter que la préservation numérique se distingue par son caractère à long terme et que la plupart des méthodes de gestion des risques sont généralement axées sur le court terme. Pour la préservation numérique, vous devez être conscient qu'à long terme, des

événements improbables deviendront plus probables et qu'il convient d'accorder une attention particulière à ceux qui ont des conséquences importantes.

Plan de continuité des activités

Raison d'être

Les parties intéressées et les parties prenantes exigent que les organisations se préparent de manière proactive à d'éventuels incidents et perturbations afin d'éviter la suspension d'opérations et de services essentiels, ou si les opérations et les services sont perturbés, qu'elles reprennent les opérations et les services aussi rapidement que le demandent ceux qui en dépendent. (ISO/PAS 22399:2007).

La planification et la pratique de la continuité des activités sont bien établies au sein de la profession informatique et ne sont pas traitées en détail dans le manuel. Cependant, elles constituent un élément important pour assurer la préservation des bits et contribuent ainsi de manière significative à la préservation numérique. Il existe une série de webinaires sur la continuité des activités et la préservation numérique dans le cadre du projet TIMBUS (voir Ressources).

L'élaboration et l'utilisation d'un plan de continuité des activités fondé sur des principes solides, approuvé par la direction générale et activé par un personnel formé, réduiront considérablement la probabilité et la gravité des conséquences des catastrophes et des incidents.

Un modèle est le plan élaboré par le *Data Archive*, et décrit dans la note de la DPC sur la continuité des activités. Les organisations peuvent également envisager d'utiliser des services dans le *cloud* (voir Services *cloud*) dans le cadre de leur planification.

Exigences

- Élaborer un plan de continuité des activités.
- Veiller à ce que tout le personnel concerné soit formé aux procédures de continuité des activités.
- Créer des copies des ressources de données au moment de leur transfert à l'institution.
- Conserver les copies sur des supports standard ou sur d'autres supports contemporains approuvés.
- Stocker les copies sur le site et hors site. Les copies hors site doivent être stockées à une distance sûre des copies sur site afin de s'assurer qu'elles ne sont pas affectées par une catastrophe naturelle ou humaine affectant les copies sur site.
- Considérer les données et les compétences comme des atouts et en dresser des registres.
- Veiller à ce que les rôles et les responsabilités soient identifiés et tenus à jour.

Ressources

ISO/PAS 22399:2007 Sécurité sociétale – Lignes directrices pour être préparé à un incident et gestion de continuité opérationnelle.
http://www.iso.org/iso/catalogue_detail?csnumber=50295
Cette norme fournit des orientations générales à toute organisation pour développer ses propres critères de performance spécifiques pour la préparation aux incidents et la continuité opérationnelle, et concevoir un système de gestion approprié.
Statut de la norme : abrogée.

ISO/IEC 27001:2013 Technologies de l'information – Techniques de sécurité – Systèmes de management de la sécurité de l'information – Exigences
http://www.iso.org/iso/catalogue_detail?csnumber=54534
Cette norme spécifie les exigences relatives à l'établissement, la mise en œuvre, la maintenance et l'amélioration continue d'un système de gestion de la sécurité de l'information. Les exigences sont génériques et sont censées être applicables à toutes les organisations.
Version française disponible à l'adresse suivante : https://www.iso.org/fr/standard/54534.html

Disaster Preparedness for Digital Content
http://dpworkshop.org/workshops/management-tools/disaster-preparedness
Un atelier en ligne sur la gestion de la préservation numérique qui relie un ensemble de 4 documents suggérés (politique de plan de catastrophe, plan de communication, plan de formation, rôles et responsabilités). Cumulativement, ils fournissent une documentation complète et sont mis à jour pour refléter les pratiques actuelles en matière de préparation aux catastrophes.

Archives nationales du Royaume-Uni, outils d'évaluation des risques
http://www.nationalarchives.gov.uk/information-management/manage-information/policy-process/digital-continuity/risk-assessment/
Les Archives nationales du Royaume-Uni fournissent deux outils d'auto-évaluation au format Excel qui sont liés à ses orientations en matière de continuité numérique et à son cadre de solutions et de services.

L'outil d'auto-évaluation (0,4 Mo) divise l'évaluation des risques en trois sections : Comprendre la continuité numérique et les rôles et responsabilités, Exigence d'information et dépendances techniques, et Gestion.

L'outil d'évaluation des risques liés aux actifs informationnels (0,26 Mo) vous aide à identifier les risques pour la continuité de tout actif d'information numérique spécifique et à déterminer où la continuité a déjà été perdue. Il formule des recommandations sur le maintien ou le rétablissement de la continuité afin de vous aider à élaborer un plan d'action pour la continuité numérique.

Boîte à outils DRAMBORA (*Digital Repository Audit Method Based on Risk Assessment*)
http://www.repositoryaudit.eu
Il s'agit d'une boîte à outils en ligne pour l'audit d'un entrepôt numérique. Il guide les utilisateurs tout au long du processus d'audit, depuis la définition de l'objectif et de la portée de l'audit jusqu'à l'identification et la gestion des risques pour l'entrepôt. DRAMBORA fournit une liste de plus de 80 exemples de risques potentiels pour les entrepôts numériques, formulés en termes de conséquences possibles.

SPOT
http://www.dlib.org/dlib/september12/vermaaten/09vermaaten.html
Le SPOT (*Simple Property-Oriented Threat*) fournit un modèle simple d'évaluation des risques, axé sur la protection contre les menaces pesant sur six propriétés des objets numériques fondamentales pour leur préservation : disponibilité, identité, persistance, capacité de rendu, compréhensibilité et authenticité. Le modèle examine les menaces en termes d'impact potentiel sur ces propriétés, en fournissant plusieurs exemples de résultats pour chacune d'entre elles. L'article décrivant le modèle comprend également une comparaison utile avec d'autres modèles de menaces pour la préservation numérique.

Managing Digital Continuity
http://www.nationalarchives.gov.uk/information-management/manage-information/policy-process/digital-continuity/

Comprend une évaluation utile des risques et de nombreuses corrélations avec les stratégies de gestion des risques pour la planification de la continuité des activités.

Step 3: Assess and manage risks to digital continuity
http://www.nationalarchives.gov.uk/information-management/manage-information/policy-process/digital-continuity/step-by-step-guidance/step-3/
Les Archives nationales du Royaume-Uni ont mis au point un outil d'auto-évaluation pour le secteur public au sens large, lié à ses orientations en matière de continuité numérique et à son cadre de solutions et de services.

Assess risks to digital continuity
http://www.nationalarchives.gov.uk/documents/information-management/assess-dc-risks-factsheet.pdf
(2 pages)

Risk Assessment Handbook
http://www.nationalarchives.gov.uk/documents/information-management/Risk-Assessment-Handbook.pdf
(35 pages)

The Atlas of Digital Damages
https://www.flickr.com/groups/2121762@N23/
Il s'agit d'une zone de mise en scène pour la collecte d'exemples visuels de problèmes de préservation numérique, de rendus ratés, de dommages d'encodage, de données corrompues et

de preuves visuelles documentant les échecs de toute nature. Vous pouvez contribuer uniquement avec une image, raconter l'histoire derrière l'image, ou partager le fichier (ou l'ensemble de fichiers) original, afin que les développeurs d'outils puissent tirer des leçons des dommages numériques et tester leur code avec eux.

Projet TIMBUS :
Business Continuity Management 1 – Intro, Life Cycle, Planning, Scope
https://www.youtube.com/watch?v=25EhtuE3XkE
1 des 4 webinaires sur la gestion de la continuité des activités et la préservation numérique des processus du projet Timbus financé par l'Union européenne. Cette introduction est probablement la plus accessible pour les novices (sortie en 2013. 13 min).

Études de cas

Business continuity procedures – UK Data Archive, University of Essex
https://www.dpconline.org/blog/1094-business-continuity-procedures-uk-data-archive-university-of-essex
Le *Data Archive* est le centre national de données pour les sciences sociales du Royaume-Uni, financé par le Conseil de recherche économique et sociale (ESRC). L'Archive est certifié ISO 27001, la norme internationale pour la sécurité de l'information, qui exige que la continuité de la sécurité de l'information soit intégrée dans les systèmes de gestion de la continuité des activités d'une organisation. Le système de stockage numérique du *Data Archive* est basé, à des fins de sécurité, sur un stockage et un accès séparés et distribués. La continuité des activités est basée sur la résilience fournie par la création de copies multiples des données et des procédures de récupération spécifiques, ainsi que sur la prévention préventive des défaillances. Chaque fichier de n'importe quel ensemble de données a au moins trois copies. L'Archive crée également une copie d'archive en lecture seule de chaque étude et de toute mise à jour au fur et à mesure qu'elle est disponible sur le système.

Références

ISO/PAS 22399:2007*Sécurité sociétale -Lignes directrices pour être préparé à un incident et gestion de continuité opérationnelle*. Disponible en ligne à l'adresse suivante : http://www.iso.org/iso/catalogue_detail?csnumber=50295

Formation et développement des compétences du personnel

Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Introduction

Une main-d'œuvre bien qualifiée et efficace peut être le plus grand atout d'une organisation, mais on n'accorde pas toujours l'attention et le soin nécessaires pour fournir une formation et un développement des compétences adéquats ou pour encourager leur acquisition. En outre, l'élaboration et le maintien d'un programme de préservation numérique peuvent sembler décourageants à bien des égards et, en particulier, lorsqu'est identifié un manque de compétences du personnel. Cela peut être dû au fait que l'environnement de travail est caractérisé par :

- un changement rapide et continu ;
- un effacement des frontières au sein des institutions et entre elles ;
- une incertitude quant à la capacité de prévoir avec confiance les tendances et les besoins futurs ;
- des rôles et des responsabilités mal définis et/ou changeants ;
- une importance accrue accordée à la collaboration et au travail d'équipe ;
- une importance accrue des responsabilités à endosser.

Une formation du personnel et des activités de développement professionnel continu (DPC) soigneusement conçues peuvent jouer un rôle clé dans la réussite de la transition d'un modèle traditionnel des bibliothèques et des services d'archives vers un modèle numérique ou hybride. Une formation et un développement appropriés des compétences peuvent grandement renforcer la confiance et les capacités des membres du personnel et réduire l'anxiété liée à l'évolution de la nature du travail dans les institutions qui mènent des actions de préservation. Une approche réfléchie de la formation et du développement des compétences (par opposition au simple fait d'envoyer les gens « suivre des cours ») est susceptible de faire une différence significative :

- en aidant le personnel à exploiter efficacement la technologie et à améliorer la qualité globale du service ;
- en améliorant le niveau individuel de satisfaction et d'engagement professionnel, permettant d'améliorer la stabilité du personnel ;
- en améliorant les perspectives stratégiques de l'organisation dans son ensemble.

Les organisations doivent adopter une approche stratégique de la formation et du développement professionnel, en considérant attentivement les compétences requises, ainsi que les rôles et responsabilités nouvelles et émergentes. La question doit être clairement abordée dans toutes les politiques, stratégies et planifications pertinentes en matière de préservation numérique, et le budget consacré aux activités de sensibilisation et de développement des compétences doit faire partie intégrante de la planification des travaux de préservation numérique.

Large éventail de compétences

La réussite des travaux de préservation numérique exige un large éventail de compétences, allant de celles spécifiques au domaine, comme la connaissance des normes de métadonnées et des cadres d'audit, à des compétences plus générales telles que la planification de projet et la gestion des risques. Par conséquent, s'assurer que tous les membres du personnel possèdent des compétences appropriées spécifiques à la préservation numérique, pour la partie du processus qui leur incombe, n'est qu'un aspect de la préparation nécessaire pour les rendre à même de maximiser le potentiel de la technologie numérique. Il est très peu probable qu'une seule personne puisse posséder toutes les compétences requises pour entreprendre l'ensemble des activités de préservation numérique, de sorte que la collaboration restera la clé du succès. Une formation adaptée peut améliorer les aptitudes et les compétences individuelles, mais peut également améliorer la compréhension des autres aptitudes et compétences requises pour la réussite d'un projet collaboratif.

Un certain nombre d'initiatives différentes se sont efforcées de clarifier les aptitudes et les compétences requises pour le travail de préservation numérique et les rôles potentiels que peut endosser le personnel à différents niveaux d'ancienneté. Des exemples sont donnés ci-après.

DPOE

Le programme DPOE (*Digital Preservation Outreach and Education*) de la Bibliothèque du Congrès a défini trois niveaux de rôles (ou stades de carrière) pour le personnel dans son modèle de formation à la préservation numérique. Ces niveaux sont les suivants :

- les directeurs : ceux qui occupent des postes de direction institutionnelle ;
- les cadres : ceux qui gèrent les programmes et les services de préservation numérique ;
- les gestionnaires : les professionnels qui travaillent sur le terrain avec des objets numériques et des solutions de préservation.

DigCurV

Le projet DigCurV a adapté le modèle à trois niveaux du DPOE pour son travail de définition des compétences de base requises pour le travail de préservation numérique. Le projet DigCurV a examiné un certain nombre de questions relatives à la formation, aux aptitudes et à la montée en compétence en matière de conservation et de préservation numériques, produisant une variété de ressources utiles, notamment une base de données des formations possibles et disponibles et un cadre de programme d'études. En décrivant les compétences de base requises à chacun des trois niveaux du modèle DPOE, à travers un ensemble de « prismes », le cadre du programme d'études DigCurV constitue une excellente ressource pour ceux qui cherchent à identifier la gamme complète des aptitudes et compétences requises pour la conservation et la préservation numériques. Plus précisément, le cadre pédagogique DigCurV peut aider les utilisateurs à décrire et à comparer les formations, à développer de nouvelles ressources en matière de formation et à cartographier les compétences et les connaissances d'un individu ou d'une équipe afin d'identifier toute lacune existante en matière de compétences.

Chaque prisme est divisé en quatre sections qui couvrent :

- les connaissances et les capacités intellectuelles ;
- les qualités personnelles ;
- la conduite professionnelle ;
- la gestion et l'assurance de la qualité.

Chacune des sections contient des sous-sections supplémentaires qui énumèrent les grandes catégories de compétences individuelles. Les catégories sont conçues pour être génériques et ont donc un large périmètre d'application, bien que des exemples spécifiques de normes ou d'outils particuliers relatifs aux compétences soient disponibles sur le site web DigCurV.

DigCCurr

Le DigCCurr (*Preserving Access to Our Digital Future: Building an International Digital Curation Curriculum*) a produit une matrice en six dimensions pour identifier et organiser les domaines à couvrir dans un programme de préservation numérique. Cette matrice des connaissances et des compétences en matière de préservation numérique est une approche alternative qui peut être particulièrement utile pour les petites organisations.

Rôles et responsabilités en matière de formation et de développement des compétences

Les rôles et les responsabilités doivent être clairement définis. Le succès des programmes de formation et de développement professionnel sera fonction du degré d'harmonisation des différents rôles et responsabilités. Il est essentiel que chacun des acteurs du processus apprécie pleinement son rôle et participe activement au processus. La liste ci-dessous est un guide des diverses responsabilités qui peuvent être exigées des différents acteurs pour assurer la création et le déploiement d'un programme de formation et de développement des compétences réussi et complet.

Rôles et responsabilités des parties prenantes
Rôles et responsabilités de l'institution • Développer une stratégie d'information qui intègre la formation en informatique à la mission globale de l'institution. • Définir, en concertation avec le personnel clé, un audit des compétences, afin de déterminer quelles sont les compétences spécifiques requises pour atteindre les objectifs de l'organisation, incluant une analyse prospective de compétences, d'activités et de responsabilités nouvelles et émergentes. • Établir un équilibre entre le recrutement de compétences spécifiques et le développement efficace des talents existants. • Fournir des ressources adéquates pour la formation et le développement des compétences. • Veiller à ce que le personnel ait accès aux équipements appropriés. • Garantir l'accès à une formation pratique et à la mise en œuvre. • Encourager le travail en réseau avec les collègues d'autres institutions. • Envisager des stratégies telles que le détachement à court terme auprès d'une institution qui peut avoir plus d'expérience dans un domaine spécifique. • Impliquer le personnel dans la conception des programmes de formation et de développement des compétences. • Faciliter une communication multidisciplinaire efficace. • Adopter une vision large de ce qui constitue la formation et le développement des compétences (c'est-à-dire la combinaison de cours formels, à la fois génériques et sur

mesure, de formations informelles au sein de l'organisation, de transfert de compétences au sein de l'organisation, de la mise en réseau, etc.).
Rôles et responsabilités des associations professionnelles • Réactivité aux besoins actuels de formation et de développement. • Capacité à travailler avec les institutions pour développer des modules de formation répondant à leurs besoins.
Rôles et responsabilités de l'individu • Capacité à supporter des changements fréquents. • Capacité à être flexible. • Capacité travailler en équipe. • Capacité à communiquer (y compris à écouter) efficacement avec les différents groupes de personnel et avec les différents niveaux hiérarchiques de l'organisation. • Capacité à poursuivre activement son développement professionnel personnel par le biais de divers mécanismes. • Capacité à partager ses compétences et son expertise. • Capacité à acquérir de nouvelles compétences. • Capacité à mettre en œuvre de nouvelles compétences.

Entreprendre un audit des compétences

Un point de départ utile pour toute organisation est de mener un audit de compétences adapté spécifiquement aux besoins d'une institution donnée. Ce processus aidera à identifier les lacunes en matière de compétences qui existent et permettra de prendre des décisions éclairées sur la formation et le développement des compétences, ainsi que de mettre en évidence le cas échéant les rôles supplémentaires qui peuvent nécessiter de nouveaux membres du personnel ou de nouvelles responsabilités (et de nouvelles fiches de poste) pour les personnes déjà en fonction. Les résultats de l'audit des compétences peuvent ensuite être utilisés pour établir une étude de cas des ressources supplémentaires qui pourraient être nécessaires. En plus d'être un excellent point de départ pour améliorer le développement des compétences du personnel, il peut également être utile d'intégrer des éléments du processus dans les processus réguliers de développement et d'évaluation professionnels et du personnel.

Le cadre du programme d'études DigCurV ou la matrice des connaissances et des compétences en matière de préservation numérique peuvent constituer un outil utile lors de la réalisation d'un audit des compétences, comme ressource pour l'évaluation comparative. Il sera nécessaire d'adapter l'audit aux pratiques et aux processus de développement des compétences du personnel de chaque organisation, mais les étapes suivantes peuvent être envisagées :

1. Identifier tous les rôles au sein de l'organisation comportant des responsabilités en matière de préservation numérique. Cette étape peut être facilitée par l'examen des processus de travail et leur mise en correspondance avec des modèles tels que le modèle de référence OAIS ou le *Curation Lifecycle Model* du *Digital Curation Centre*.
2. Faire correspondre les rôles au prisme pertinent du cadre DigCurV.
3. Travailler avec les détenteurs de rôles pour faire correspondre les compétences au prisme pertinent. Cela peut se faire de différentes manières, notamment par l'auto-évaluation et dans le cadre d'une activité de groupe. Il peut également être utile de noter selon un barème.
4. Analyser les résultats pour identifier les lacunes, les besoins de formation et les rôles supplémentaires nécessaires.

Options de formation et de développement des compétences

Le manque d'opportunités de formation et de développement des compétences était auparavant un obstacle considérable pour ceux qui souhaitaient apprendre davantage sur la préservation numérique. Alors que les personnes plus avancées dans leur montée en compétences peuvent encore avoir du mal à trouver des opportunités appropriées, il existe maintenant un certain nombre de cours bien établis, accessibles pour les débutants et les niveaux intermédiaires, depuis les cours de courte durée jusqu'aux programmes complets diplômants, en passant par une variété d'opportunités de formation abordant des domaines d'intérêt spécialisés spécifiques. Le temps et les frais liés à la participation à une formation en présentiel constituent désormais un obstacle plus important, mais de plus en plus d'options de formations en ligne et à distance sont disponibles, diminuant d'autant plus cet obstacle.

Les cours sur la préservation numérique ont également souffert auparavant de critiques concernant l'accent mis sur la théorie plutôt que sur la pratique. Ceci est également en train de changer avec l'intégration dans la formation d'un plus grand nombre d'exercices pratiques et de démonstrations d'outils. La préservation numérique reste également une discipline où l'on peut apprendre autant, sinon plus, en faisant, de sorte que l'apprentissage entre pairs et la volonté de mettre les mains dans le cambouis peuvent souvent donner les meilleurs résultats. Le partage d'informations et les échanges avec des équipes d'organisations similaires peuvent constituer une méthode particulièrement efficace pour le développement des compétences et la montée en compétence des agents.

Ressources

2014 *DPOE Training NeedsAssessment Survey*[23]
https://www.digitalpreservation.gov/education/documents/DPOENeedsAssessmentSurveyExecutiveSummary.pdf
Une analyse de l'état des pratiques en matière de préservation numérique et de la capacité à préserver les contenus numériques au sein des organisations aux États-Unis dans le but d'établir les lacunes et les besoins en matière de formation (13 pages).

DigCurV, *A Curriculum Framework for Digital Curation*
http://www.digcurv.gla.ac.uk/
Le cadre du programme d'études DigCurV offre un moyen d'identifier, d'évaluer et de planifier la formation afin de répondre aux besoins en compétences du personnel engagé dans la préservation numérique. L'équipe DigCurV a entrepris une recherche multinationale dans le secteur du patrimoine culturel afin de comprendre les compétences utilisées par les personnes travaillant

[23]Enquête d'évaluation des besoins de formation publiée par le *Digital Preservation Outreach and Education Program* en 2014.

dans le domaine de la préservation numérique et celles recherchées par les employeurs dans ce secteur. Le cadre définit des objectifs de compétences propres pour répondre aux besoins spécifiques de trois publics distincts : les directeurs, les cadres et les gestionnaires.

- Les compétences définies sous le **prisme du directeur** permettent à un professionnel de la préservation numérique de maintenir une vision stratégique.
- Les compétences définies sous le **prisme du cadre** permettent à un professionnel de planifier et de suivre l'exécution de projets de préservation numérique, de recruter et de soutenir des équipes de projet, et d'assurer la liaison avec une série de contacts internes et externes au sein du secteur du patrimoine culturel.
- Les compétences définies sous le **prisme des gestionnaires** permettent à un professionnel de planifier et d'exécuter diverses tâches techniques, tant individuellement qu'au sein d'une équipe pluridisciplinaire.

Matrix of Digital Curation Knowledge and Competencies
http://ils.unc.edu/digccurr/digccurr-matrix.html
Le DigCCurr(*Preserving Access to Our Digital Future: Building an International Digital Curation Curriculum*) a produit une matrice en six dimensions pour identifier et organiser les domaines à couvrir dans un programme d'études de préservation numérique.

Digital Preservation Outreach and Education
http://www.digitalpreservation.gov/education/curriculum.html
Le programme de formation de base de la Bibliothèque du Congrès sur la préservation numérique pour le personnel de gestion des archives et des collections. Le cours complet est dispensé aux archivistes et aux autres professionnels de la préservation numérique selon une approche de « formation des formateurs », afin de favoriser une diffusion plus large auprès des collègues. Les vidéos de présentation sont disponibles en ligne.

Les formations proposées par *Digital Preservation Coalition*
https://www.dpconline.org/digipres/train-your-staff
Un des rôles clés de la DPC est de responsabiliser et de développer les compétences de ses membres en activité. La DPC s'occupe de cette question en facilitant les activités de formation et de soutien et en élaborant des livrables et organisant des événements centrés sur les besoins du personnel opérationnel tout au long de l'année. Parmi ceux-ci figurent *The DPC Leadership Programme*, *The Digital Preservation Roadshow*, *The Member Briefing Days*et les évènements sur invitation.

Le programme de formation à la préservation numérique[24]
https://dptp.london.ac.uk/
Une série de cours de formation à la préservation numérique basés au Royaume-Uni. Les cours DPTP programmés s'étendent sur deux ou trois jours ours et ont lieu régulièrement tout au long de l'année.

Digital Preservation Management: Implementing short-term strategies for long-terms problems
http://dpworkshop.org/
Un excellent tutoriel en ligne gratuit qui vous présente les principes de base de la préservation numérique. Il est particulièrement destiné aux bibliothécaires, archivistes, conservateurs, cadres et

[24] *Digital Preservation Training Programme* (DPTP).

spécialistes techniques. Il comprend des définitions, des concepts clés, des conseils pratiques, des exercices et des références actualisées. Le tutoriel est disponible en anglais, français et italien.

Cours universitaires de troisième cycle au Royaume-Uni

https://www.dpconline.org/digipres/train-your-staff

La DPC tient une liste qui sera utile à tous ceux qui envisagent d'obtenir un diplôme de troisième cycle centré sur la préservation numérique. Elle comprend des options d'enseignement universitaire dispensées sur le campus et à distance. Certaines universités proposent également des modules individuels donnant droit à des crédits sur des sujets pertinents en matière de préservation numérique.

Éducation et formation en matière d'archivage et de préservation des objets audiovisuels

http://www.arsc-audio.org/etresources.html

Des possibilités de formation en Australie, en Europe et aux États-Unis pour ceux qui travaillent sur des contenus sonores et des images mobiles.

Connecting to Collections: Caring for Audiovisual Material

http://www.connectingtocollections.org/av/

Des cours à rythme libre comprenant des webinaires enregistrés, des documents, des diaporamas et des suggestions de lectures complémentaires pour chaque étudiant travaillant sur des objets audiovisuels. Il couvre les principes de base, l'histoire des formats et les défis posés par leur préservation, l'identification des formats, les questions d'accès et un aperçu des modèles et normes existants. Il est rédigé en anglais par une équipe d'archivistes, de conservateurs et d'experts en préservation numérique basés aux États-Unis.

How the DPC makes a difference to your staff

https://vimeo.com/45433968

De courts entretiens avec cinq candidats parrainés par *Digital Preservation Coalition* pour participer à la *Digital Futures Academy* à Londres en mars 2012. Ils réfléchissent à leur expérience et à la manière dont leur adhésion à la DPC a profité à leurs institutions (2 minutes 40 secondes).

Normes et bonnes pratiques

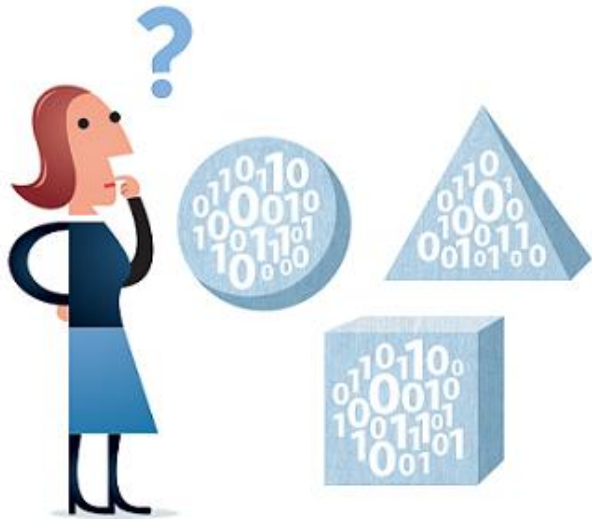


Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Introduction

L'utilisation et le développement de normes fiables ont été depuis longtemps une pierre angulaire de l'industrie de l'information. Elles facilitent l'accès, la découverte et le partage des ressources numériques, ainsi que leur préservation à long terme. Il existe à la fois des normes génériques applicables à tous les secteurs, qui peuvent soutenir la préservation numérique, et des normes spécifiques à un secteur qu'il peut être nécessaire de respecter. L'utilisation de normes adaptées à l'environnement institutionnel numérique contribue à la conformité organisationnelle et à l'interopérabilité entre les systèmes divers au sein et en dehors du secteur. Le respect des normes permet également aux organisations d'être auditées et certifiées.

Normes opérationnelles

Il existe un certain nombre de normes qui peuvent aider à l'élaboration d'un modèle opérationnel de préservation numérique.
Prendre le contrôle de la préservation des objets numériques nécessite un ensemble de procédures pour régir leur transfert dans un environnement de préservation numérique. Cela peut inclure l'identification et l'estimation volumétrique des objets à transférer, l'évaluation des coûts de préservation associés et l'identification des exigences d'authentification et de confidentialité futures.

La norme ISO 20652:2006 – Systèmes de transfert de données et d'informations spatiales – Interface entre producteur et Archives – Norme pour la méthodologie concernant les abstraits (ISO, 2006) est une norme internationale qui fournit un cadre méthodologique pour le développement de procédures en matière de transfert formel d'objets numériques depuis l'environnement du producteur dans l'environnement de préservation numérique. Les objectifs, les actions et les résultats attendus sont identifiés par quatre phases - les négociations initiales avec le producteur (phase préliminaire), la définition des exigences (phase de définition formelle), le transfert des objets numériques dans l'environnement de préservation numérique (phase de transfert) et la garantie que les objets numériques et les métadonnées qui les accompagnent sont conformes à ce qui a été convenu (phase de validation).

La norme ISO 14721:2012 – Systèmes de transfert de données et d'informations spatiales - Système ouvert d'archivage d'information (SOAI) – Modèle de référence (OAIS) (ISO, 2012b) fournit un cadre systématique pour comprendre et mettre en œuvre les concepts archivistiques nécessaires à la préservation de l'information numérique et à son accès à long terme, et pour décrire et comparer les architectures et les opérations des archives existantes et futures. Elle décrit les rôles, les processus et les méthodes pour la préservation à long terme. Élaborée par le Comité consultatif pour les systèmes de données spatiales (CCSDS), la norme OAIS a été publiée pour la première fois en 1999 et a eu une influence sur de nombreuses initiatives en matière de préservation numérique depuis le début des années 2000. Un guide utile introductif à la norme est disponible sous la forme d'un rapport de veille technologique de la DPC (Lavoie, 2014).

La norme OAIS est « une Archive, consistant en une organisation de personnes et de systèmes qui a accepté la responsabilité de pérenniser des informations et de les rendre accessibles pour une communauté désignée définie ». Une « Archive OAIS » peut être distinguée des autres utilisations du terme « archive » dans la mesure où elle accepte et répond à une série de responsabilités spécifiques. L'OAIS définit ces responsabilités comme suit :

- négocier et accepter des informations appropriées de la part des producteurs d'informations ;
- obtenir un contrôle suffisant de l'information afin d'atteindre les objectifs de préservation à long terme ;
- -déterminer l'étendue de la communauté des utilisateurs des archives ;
- s'assurer que l'information pérennisée est compréhensible de manière indépendante par la communauté des utilisateurs, en ce sens que l'information peut être comprise par la communauté des utilisateurs sans l'aide du producteur de l'information ;
- suivre des politiques et des procédures documentées pour garantir que l'information est pérennisée contre toute éventualité raisonnable et qu'il n'y a pas de suppressions ad hoc ;
- mettre l'information pérennisée à la disposition de la communauté des utilisateurs et permettre la diffusion de copies authentifiées de l'information pérennisée sous sa forme originale ou sous une forme permettant de remonter jusqu'à l'original (Lavoie, 2014).

La norme OAIS définit également le modèle d'information qui doit être adopté. Celui-ci comprend non seulement les objets numériques, mais aussi toutes les métadonnées utilisées pour décrire ou gérer l'objet numérique et toute autre information d'appui appelée « Information de représentation ».

Le modèle fonctionnel OAIS est largement utilisé pour établir des processus de travail et des mises en œuvre techniques. Il définit un large éventail de fonctions de préservation numérique, notamment le versement, l'accès, le stockage, la planification de la préservation, la gestion et l'administration des données. Ces fonctions fournissent un ensemble commun de concepts et de définitions qui peuvent faciliter les discussions entre les secteurs et les groupes professionnels et faciliter la spécification des archives et des systèmes de préservation numérique.

La norme OAIS fournit un cadre de haut niveau et un langage partagé utile pour la préservation numérique, mais, pendant de nombreuses années, le concept de la « conformité à l'OAIS » est resté difficile à cerner. Bien que le terme ait été fréquemment utilisé dans les années suivant immédiatement la publication de la norme, il reposait sur la capacité de mesurer jusqu'à six responsabilités obligatoires mais de haut niveau. Une discussion plus détaillée sur la "conformité à l'OAIS" se trouve dans le rapport de veille technologique.2

La norme ISO/TR 18492:2005 – *Long-term preservation of electronic document-based information* (ISO/TR, 2005) fournit une méthodologie pratique pour la préservation et la récupération continues d'informations basées sur des documents électroniques authentiques, qui comprend des conseils technologiquement neutres sur le renouvellement des supports, la migration, la qualité, la sécurité et le contrôle environnemental. Ces conseils sont élaborés pour garantir l'authenticité des documents au-delà de la durée de vie des systèmes qui conservent les informations d'origine.

La norme ISO 15489:2001 – *Information et documentation : « Records management* » (ISO, 2001) peut également être une norme utile pour définir les rôles, les processus et les méthodes pour une mise en œuvre de la préservation numérique où l'accent est mis sur la gestion à long terme des documents. Cette norme définit un cadre de bonnes pratiques pour la gestion des documents d'activités afin de garantir qu'ils sont conservés et documentés tout au long de leur cycle de vie tout en restant fiables et accessibles. Une version révisée a été publiée en 2016 (ISO, 2016).

La norme ISO 16175:2011 – *Principes et exigences fonctionnelles pour les enregistrements dans les environnements électroniques de bureau* (ISO, 2011) concerne les systèmes de gestion des documents électroniques et des documents d'activité ainsi que les systèmes de gestion de contenus d'entreprise. Bien qu'elle ne comporte pas d'exigences spécifiques pour la préservation numérique, elle reconnaît la nécessité de conserver les documents au fil du temps et que les questions d'obsolescence des formats doivent être prises en compte dans la spécification de ces systèmes électroniques. Une version révisée de la norme a été publiée en 2020 (ISO, 2020).

Il existe des normes internationales génériques qui peuvent également être pertinentes dans le domaine de la préservation numérique.

- La certification ISO 9001 – *Systèmes de management de la qualité* (ISO, 2015) démontre la capacité d'une organisation à fournir et à améliorer des produits et des services cohérents ;
- La certification ISO/IEC 27001 – *Technologies de l'information -Techniques de sécurité – Systèmes de gestion de la sécurité de l'information* (ISO/IEC, 2013) démontre que les objets numériques sont gérés de manière sûre en garantissant leur authenticité, leur fiabilité et leur facilité d'utilisation ;
- La norme ISO/IEC 15408 – *Critères communs pour l'évaluation de la sécurité des technologies de l'information* (ISO/IEC, 2009) fournit un cadre normalisé pour la spécification des exigences fonctionnelles et en matière d'assurance pour la sécurité des TI, ainsi qu'une évaluation rigoureuse de celles-ci.

Il existe un certain nombre de voies par lesquelles une mise en œuvre de la préservation numérique peut être certifiée. Celles-ci vont des méthodes de certification par examen collégial léger, telles que le *Data Seal of Approval*[25], aux méthodes internes plus étendues de la norme DIN 31644 *Information et documentation– Critères pour les archives numériques fiables* (DIN, 2012), en passant par la norme internationale complète ISO 16363:2012 – *Audit et certification de référentiels numériques de confiance* (ISO, 2012a) (voir Audit et certification).

[25]Sceau d'approbation des données.

Normes techniques

L'utilisation de normes pour les aspects techniques d'un programme de préservation numérique présente des avantages spécifiques, principalement en ce qui concerne les métadonnées et les formats de fichiers.

Conjointement avec les normes de métadonnées descriptives pertinentes, PREMIS et METS sont des standards qui amélioreront *de facto* un programme de préservation numérique. PREMIS (*PREservation Metadata: Implementation Strategies*)[26] est un standard hébergé par la Bibliothèque du Congrès et publiée pour la première fois en 2005. Le dictionnaire de données et les outils de soutien ont été spécifiquement conçus pour soutenir la préservation de l'objet numérique. METS (*Metadata Encoding and Transmission Standard*) [27]est un standard d'encodage XML qui permet d'empaqueter les objets numériques avec des informations archivistiques (voir Métadonnées et documentation).

Il existe également des normes relatives aux formats de fichier. Le choix de formats de fichier non propriétaires et basés sur des normes de format ouvert donne à une organisation une bonne base pour un programme de préservation numérique. La norme ISO/IEC 26300-1:2015 – *Format de document ouvert pour applications de bureau* (ISO/IEC, 2015) fournit un schéma XML pour la préservation des documents largement utilisés tels que les documents texte, les feuilles de calcul, les présentations. La norme ISO 19005 –*Format de fichier de document électronique pour une préservation à long terme* (ISO, 2005) prescrit les éléments d'un PDF/A valide qui garantit que les fichiers sont autoporteurs et s'affichent correctement sur différents appareils. Certains aspects des formats JPEG-2000 et TIFF sont également couverts par des normes ISO (voir Formats de fichiers et normes).

Obstacles à l'utilisation des normes

Une approche de la préservation numérique basée sur des normes est importante, mais il existe également des facteurs qui empêchent leur utilisation en tant que stratégie de préservation numérique :

- Le rythme du changement est si rapide que les normes qui ont atteint le stade de l'approbation officielle, un processus qui prend généralement des années, seront inévitablement à la traîne et pourraient même être dépassées.
- Les pressions concurrentielles entre fournisseurs encouragent le développement d'extensions propriétaires ou la mise en œuvre de normes qui peuvent diluer les avantages de la cohérence et de l'interopérabilité pour la préservation.
- Les normes elles-mêmes s'adaptent et changent en fonction des nouveaux environnements technologiques, ce qui entraîne un certain nombre de variations de la norme originale qui peuvent ou non être interopérables à long terme, même si elles sont rétro compatibles à court terme.
- Les normes peuvent être intimidantes à lire et leur mise en œuvre peut nécessiter beaucoup de ressources.
- Dans un tel environnement changeant et hautement distribué, il est impossible d'être complètement normatif.

Ces facteurs signifient que les normes devront être considérées comme faisant partie d'un ensemble de stratégies de préservation plutôt que comme la stratégie clé elle-même.

[26] Préservation des métadonnées : stratégies de mise en œuvre.
[27] Norme d'encodage et de transmission des métadonnées.

L'environnement numérique n'est pas enclin à être contraint par des règles rigides et un programme de préservation numérique peut souvent être un mélange de normes et des meilleures pratiques, suffisamment souple et adapté pour répondre aux besoins de l'organisation, à sa situation et aux objets numériques qu'elle gère.

Normes, meilleures pratiques et bonnes pratiques

Ces dernières années, les Archives nationales, les Bibliothèques nationales et d'autres organisations culturelles ont publié des guides de bonnes pratiques et des études de cas. La préservation numérique est également un sujet très discuté sur les blogs et les médias sociaux qui peuvent souvent fournir des informations en temps réel en relation avec la théorie et la pratique du monde entier. Les articles présentés lors de conférences telles que l'iPRES, *l'International Digital Curation Conference* (*IDCC*) et le *Preservation and Archiving Special Interest Group* (*PASIG*) peuvent être une source utile de réflexion actualisée émanant des universitaires et des professionnels de la préservation numérique.

Les normes doivent être comprises comme une description et une reconnaissance formelles de ce qu'une communauté d'experts pourrait appeler les meilleures pratiques. Les normes et les meilleures pratiques dont elles découlent peuvent être intimidantes et il existe un risque pour ceux qui se lancent dans la préservation numérique que le « mieux devienne l'ennemi du bien ». Ainsi, en adoptant ou en recommandant des normes, il faut toujours comprendre qu'il vaut presque toujours mieux agir que ne rien faire. La préservation numérique est une activité désordonnée qui engendre des défis inattendus. Il est donc presque toujours préférable d'adopter une norme mal appliquée que d'attendre la perfection.

Exigences sectorielles spécifiques

Des secteurs d'activités spécifiques sont devenus actifs dans l'élaboration de normes de préservation, et des types particuliers de contenus et de cas d'utilisation sont apparus qui chevauchent et prolongent un certain nombre de normes. Il est très avantageux que les normes de préservation numérique soient intégrées dans des normes sectorielles, car cela facilitera grandement leur adoption, même si cela peut poser un problème de coordination des activités. Trois exemples sont donnés ci-dessous :

1. Les objets audiovisuels constituent un cas particulier de préservation numérique (voir Images animées et son). Des recommandations pour les enregistrements audio et vidéo existent sous les auspices de l'*International Association of Sound and Audio-visual Archives* (telles que l'IASA- TC04, 2009), tandis qu'une série d'entreprises industrielles et de détenteurs de contenu, dont la *BBC*, la RAI, l'ORF et l'INA, ont formé le *PrestoCentre* pour faire progresser la recherche et le développement de normes de préservation dans ce domaine. https://www.prestocentre.org/[28]
2. L'industrie aérospatiale a des exigences particulières en matière de gestion du cycle de vie des produits et d'échange d'informations qui ont donné lieu à une série d'initiatives à l'échelle de l'industrie visant à normaliser les approches en matière d'alignement et de partage des dessins CAO pour l'ingénierie. L'organisme PROSTEP a créé la norme ISO 10303–*Norme pour l'échange de données modèles de produits* qui est devenue la norme LOTAR (http://www.lotar-international.org/lotar-standard). La norme LOTAR n'est pas

[28] NdT : lien cassé, archive disponible sur : https://web.archive.org/web/20190506092542/https://www.prestocentre.org/.

incompatible avec la norme OAIS, mais comme elle s'inscrit dans un protocole d'échange de données important pour l'industrie, les ingénieurs aérospatiaux sont plus susceptibles de rencontrer la norme LOTAR que la norme OAIS.

3. L'association *The Storage Network Industry Association* a également commencé à progresser dans l'élaboration d'une série de normes. Un groupe de travail de la SNIA sur la préservation des données à long terme est responsable de la préservation physique et logique, ainsi que de la création d'architectures, de services et d'interfaces de référence pour la préservation. En outre, un groupe de travail sur le stockage dans le *cloud* devrait devenir particulièrement influent en ce qui concerne la préservation. Les architectures *cloud* modifient la façon dont les organisations considèrent les entrepôts et la façon dont elles accèdent aux services pour les gérer. Par exemple, on ne sait pas très bien comment on pourrait mesurer le succès d'un « entrepôt numérique fiable » qui serait basé sur un fournisseur de services *cloud*.

Ressources

Seeing Standards: A Visualization of the Metadata Universe
http://jennriley.com/metadatamap/
Le nombre de normes de métadonnées dans le secteur du patrimoine culturel est énorme, et leurs interrelations compliquent encore la situation. Cette carte visuelle du paysage des métadonnées est destinée à aider les planificateurs dans la sélection et la mise en œuvre des normes de métadonnées. Chacune des 105 normes énumérées ici est évaluée en fonction de son degré d'application à des catégories définies dans chacun des quatre axes : communauté, domaine, fonction et objectif (2010, 1 page).

Dlib Magazine
http://www.dlib.org/dlib.html
Le *Dlib Magazine* publie régulièrement un large éventail d'articles et d'études de cas sur la mise en œuvre pratique des normes et des meilleures pratiques en matière de préservation numérique.

Core Trust Seal
https://www.coretrustseal.org/

PREMIS
http://www.loc.gov/standards/premis/
Bibliothèque du Congrès, 2015.

Version française : Priscilla Caplan*, Comprendre PREMIS, Library of Congress Network Development and MARC Standards Office*, 2009. Traduction par Louise Fauduet, Clément Oury, Sébastien Peyrard, Jean-Philippe Tramoni, Bibliothèque nationale de France.
https://www.loc.gov/standards/premis/Understanding-PREMIS_french.pdf

The Digital Curation Centre

http://www.dcc.ac.uk/

The *Digital Curation Centre* met à disposition des recherches et des études de cas en rapport avec la préservation des données de recherche. Il publie également des enregistrements des actes de sa conférence internationale annuelle sur la préservation numérique.

The Signal

http://blogs.loc.gov/digitalpreservation/

The Signal est un blog sur la préservation numérique publié par la Bibliothèque du Congrès.

IPRES

http://www.ipres-conference.org/

IPRES, la Conférence internationale sur la préservation numérique publie un site web et les actes de son événement annuel qui examine différents thèmes au sein du paysage de la préservation numérique.

The Digital Preservation Coalition Wiki

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

Le wiki de la *Digital Preservation Coalition* met à disposition un espace de collaboration pour les utilisateurs de la norme OAIS, et les évaluations de formats de fichier de la *British Library* ainsi que d'autres ressources.

Digital Preservation Matters

http://preservationmatters.blogspot.co.uk/

Le blog *Digital Preservation Matters* est un compte personnel exposant des expériences de travail sur la préservation numérique

Références

DIN, 2012. *DIN 31644 Information et documentation - Critères pour les archives numériques fiables (indisponible en français)*. Disponible à l'adresse suivante : http://data-archive.ac.uk/curate/trusted-digital-repositories/standards-of-trust?index=3

IASA-TC04, 2009.*Guidelines in the Production and Preservation of Digital Audio Objects: standards, recommended practices, and strateg*[ie]*s: 2nd edition,* edited by Kevin Bradley. Disponible à l'adresse suivante : http://www.iasa-web.org/tc04/publication-information

Une version française de la seconde édition a été publiée en 2015 (cf. le site de l'IASA : https://www.iasa-web.org/iasa-special-and-technical-publications)

ISO, 2001. *ISO 15489:2001 – Information et documentation – Records management*. Geneva : International Organization for Standardization. Disponible à l'adresse suivante : http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=31908

Norme révisée par l'ISO 15489-1:2016 – *Information et documentation – Gestion des documents d'activité - Partie 1 : Concepts et principes (NdT).* Disponible à l'adresse suivante : https://www.iso.org/fr/standard/62542.html

ISO, 2005. *ISO 19005-1:2005 – Gestion des documents – Format de fichier des documents électroniques pour une conservation à long terme (indisponible en français)*. Geneva: International Organization for Standardization. Disponible à l'adresse suivante : http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=38920

ISO, 2006. *ISO 20652:2006 – Systèmes de transfert des informations et données spatiales– Interface entre producteur et archives – Norme pour la méthodologie concernant les abstraits (indisponible en français)*. Geneva: *International Organization for Standardization*. Disponible à l'adresse suivante : http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=39577

ISO, 2011. *ISO 16175 2:2011 – Principes et exigences fonctionnelles pour les enregistrements dans les environnements électroniques de bureau – Partie 2 : Lignes directrices et exigences fonctionnelles pour les systèmes de management des enregistrements numériques.* Geneva : *International Organization for Standardization*. Disponible à l'adresse suivante : http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=55791

Norme révisée par l'ISO /TS 16175-2:2020–*Information et documentation – Processus et exigences fonctionnelles applicables aux logiciels de gestion des documents d'activité – Partie 2 : Recommandations pour le choix, la conception, la mise en œuvre et la maintenance des logiciels gérant des documents d'activité.* Disponible à l'adresse suivante : https://www.iso.org/fr/standard/74293.html

ISO, 2012a. *ISO 16363:2012 – Audit et certification des référentiels numériques de confiance (indisponible en français).* Geneva : *International Organization for Standardization*. Disponible à l'adresse suivante : http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=56510

ISO, 2012b. *ISO 14721:2012 – Systèmes de transfert des informations et données spat–ales - Système ouvert d'archivage d'information (OAIS).* Geneva : *International Organization for Standardization*. Disponible à l'adresse suivante : http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=57284

Version française : Modèle de référence pour un Système ouvert d'archivage d'information (OAIS), Pratique recommandée, CCSDS 650.0-M-2 (F), Livre Magenta, juin 2012 (version française d'octobre 2017). Disponible à l'adresse suivante : https://public.ccsds.org/Pubs/650x0m2%28F%29.pdf

ISO, 2015. *ISO 9001:2015 – Systèmes de management de la qualité - Exigences. Geneva : International Organization for Standardization*. Disponible à l'adresse suivante : http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=62085

Version française disponible à l'adresse suivante :https://www.iso.org/fr/standard/62085.html

ISO/IEC, 2009. *ISO/IEC –15408:2009 Technologies de l'information - Techniques de sécurité - Critères d'évaluation pour la sécurité TI – Partie 1 : Introduction et modèle général*. Geneva : *International Organization for Standardization*. Disponible à l'adresse suivante :

http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=50341

Version française disponible à l'adresse suivante : https://www.iso.org/fr/standard/50341.html

ISO/IEC, 2013. *ISO/IEC 27001:2013 Technologies de l'information – Techniques de sécurité– Systèmes de management de la sécurité de l'information – Exigences*. Geneva : International *Organization for Standardization*. Disponible à l'adresse suivante :
http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=54534

Version française disponible à l'adresse suivante :https://www.iso.org/fr/standard/54534.html

ISO/IEC, 2015. *ISO/IEC 26300-1:2015 – Technologies de l'information–Format de document ouvert pour applications de bureau (OpenDocument) v. 1.2 – Partie 1: Schéma de document ouvert (indisponible en français)*. Geneva : *International Organization for Standardization*. Disponible à l'adresse suivante :
http://www.iso.org/iso/home/store/catalogue_ics/catalogue_detail_ics.htm?csnumber=66363

ISO/TR, 2005.*ISO/TR 18492:2005 –Long-term preservation of electronic document-based information (indisponible en français)*. Geneva : *International Organization for Standardization*. Disponible à l'adresse suivante :
http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=38716

Lavoie, B., 2014. *The Open Archival Information System (OAIS) Reference Model: Introductory Guide(2nd Edition). DPC Technology Watch Report 14-02*. Disponible à l'adresse suivante :
http://dx.doi.org/10.7207/twr14-02

Bibliothèque du Congrès, 2015. *METS Metadata Encoding and Transmission Standard*. Disponible à l'adresse suivante : http://www.loc.gov/standards/mets/

Analyse de rentabilité, avantages, coûts et impact

1

Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Introduction

Tout changement dans l'environnement économique peut signifier pour de nombreuses organisations une réduction de leurs dépenses globales et une optimisation de leur efficacité. En parallèle, les organisations conservent des volumes croissants d'objets numériques. La réutilisation de modèles peut constituer une partie de la réponse à ce défi. La gestion à long terme, la préservation, des objets numériques est une activité coûteuse et complexe. Elle ne peut être réalisée de manière fiable sans engager des ressources et des dépenses.

Le défi pour une organisation est de créer des modèles de gestion qui :

- peuvent aider à définir des avantages et des résultats, et convaincre les principaux décideurs que la démarche en vaut la peine ;
- soutiennent les buts et objectifs plus larges de l'organisation mère ;
- assurent l'avenir de leurs objets numériques d'une manière durable et aident à concilier la différence entre les engagements de financement/les cycles de financement à court et moyen terme, et les objectifs de préservation à long terme.

D'autres organisations ont déjà créé des études de cas et des modèles de calcul des coûts et des avantages. Ainsi, la réutilisation de certains modèles ou de certaines parties de ces modèles peut non seulement permettre de gagner du temps, mais aussi servir à justifier l'adoption de stratégies particulières.

Études de cas

L'étude de cas est un outil qui permet de défendre et de garantir qu'un investissement est justifié en termes d'orientation stratégique de l'organisation et des avantages qu'il apportera. Elle présente généralement un contexte, des avantages, des coûts et un ensemble d'options aux principaux décideurs et financeurs. Elle peut également indiquer comment le succès sera mesuré afin de garantir que les améliorations promises soient réalisées.

Il est essentiel que tout modèle de gestion ou proposition qui est créé soutienne les buts et objectifs plus larges de l'organisation mère. Il est tout aussi important que les principales parties prenantes, telles que les responsables des budgets, soient consultées, qu'elles aient un aperçu préliminaire des opérations planifiées et qu'elles aient la possibilité de faire des commentaires et d'apporter leur contribution. Une exposition précoce des opérations planifiées peut, dans une

certaine mesure, améliorer des situations dans lesquelles ces opérations pourraient autrement être rejetées d'emblée.
Toutefois, la présentation à un stade précoce d'une étude de cas pour la préservation ne garantit pas qu'elle sera acceptée. Bien qu'il n'existe pas de modèle infaillible, certaines ou toutes les étapes suivantes peuvent être utiles si un plan est rejeté. Au sein d'une organisation, il peut exister des procédures et des politiques concernant l'élaboration et la présentation d'études de cas et celles-ci doivent être suivies. Une communication précoce du plan d'activités peut aider à identifier les sujets ou les domaines qui pourraient poser problème lors de la présentation officielle du plan.

Identifier les options et être pragmatique

Le but de la planification des activités est d'être ambitieux et de créer des services ou des produits qui ont une valeur et des avantages. Les avantages de la préservation à long terme ne sont pas toujours clairs, surtout lorsque les coûts sont un problème récurrent ou lorsque des ressources doivent être engagées à long terme. La planification des activités est souvent un exercice pratique. Il peut être plus efficace de faire un certain nombre de petits plans d'activités plus ciblés qu'une seule grande proposition. Grâce à sa connaissance d'une organisation, l'auteur d'un plan d'activités doit s'assurer que ce dernier est réaliste et cohérent avec les moyens de l'organisation. La planification stratégique fournit le cadre dans lequel les plans d'activités sont rédigés. Tout objectif stratégique peut être atteint de plusieurs manières, par exemple en dépensant moins d'argent mais en consacrant plus de temps, en mobilisant moins de personnel mais en allongeant le calendrier, etc. Une réponse pragmatique offre aux décideurs une option privilégiée et les raisons de cette préférence, ainsi qu'une petite gamme d'options alternatives dans le cadre de l'étude de cas. Il est souvent utile d'inclure les « coûts/bénéfices de l'inaction » comme une option par rapport à laquelle d'autres actions peuvent être évaluées.

Rejet préliminaire

Travaillez avec les parties prenantes pour identifier les raisons pour lesquelles un plan d'activités a été rejeté. Parlez avec les personnes impliquées dans la prise de décision et demandez un retour d'information spécifique. Le coût était-il trop élevé ? Les plans étaient-ils trop ambitieux ? L'étude de cas était-elle mal rédigée ou mal présentée ? Le calendrier ne correspondait-il pas aux plans de l'organisation ?
Réponse : travaillez avec les parties prenantes pour répondre aux principales préoccupations. Soyez clair pour aborder chaque question. Expliquez les raisons pour lesquelles ce plan d'activité a été présenté et ce qu'il vise à réaliser. Concentrez-vous sur les avantages, en particulier ceux qui concernent les principaux objectifs stratégiques de votre organisation. Concentrez-vous aussi bien sur les avantages à court terme qu'à plus long terme du plan d'activités. Une bonne approche consiste à créer des plans d'activités « SMART[29] », c'est-à-dire spécifiques, mesurables, réalisables, réalistes et opportuns.

Changement de la situation

Le plus difficile en matière de planification d'activités est d'arriver au point où un plan est accepté. Cependant, les circonstances peuvent changer. Si un plan d'activités n'est pas mis en œuvre ou si un financement précédemment convenu est retiré, les conséquences peuvent être graves. Là encore, la communication avec les principales parties prenantes est essentielle et peut révéler les causes du changement.

[29] De l'anglais, SMART : Specific, Measurable, Achievable, Realistic, Timely.

Réponse : une partie de la planification des activités consiste à disposer d'un éventail d'options pouvant être proposées en cas de problèmes liés au financement d'une option privilégiée. Disposer d'un plan d'activités bien structuré dont certaines propositions peuvent être supprimées peut aider à présenter un autre argument en faveur d'une mise en œuvre progressive ou alternative nécessitant moins de ressources. Dans un tel cas, le plan d'activités pourrait être rapidement remanié dans des termes plus acceptables et des ressources pourraient être mises à disposition. Mettre l'accent sur les raisons pour lesquelles les ressources n'ont pas été mises à disposition permet de présenter à nouveau un cas d'étude en soulignant davantage les avantages et l'impact positif.

Élaborer des études de cas

Les étapes suivantes doivent être prises en compte lors de la rédaction et de la présentation d'une étude de cas.

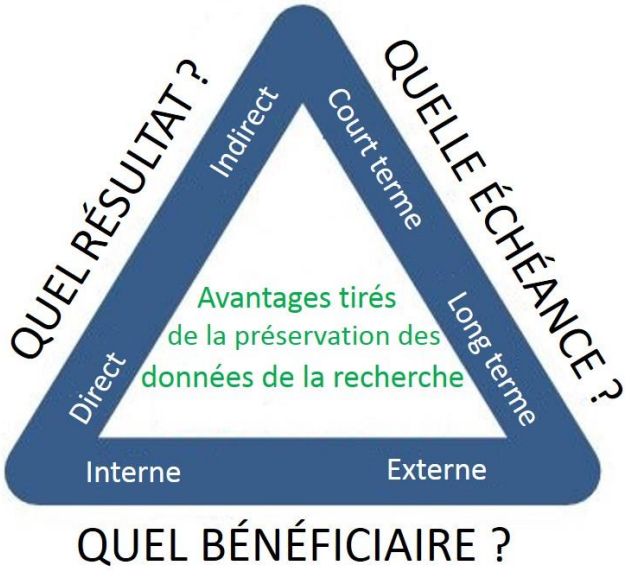
Élaborer une étude de cas	
1. Vérifier vos objets numériques et fixer un ordre de priorité des tâches requises	Faites un audit de vos objets numériques. Analysez les risques et les opportunités pour vos objets numériques. Utilisez votre analyse pour classer les domaines de travail par ordre de priorité et leur attribuer des responsables.
2. Est-ce le bon moment ?	Quelles sont vos activités actuelles ? Est-ce le bon moment pour entreprendre de nouvelles initiatives par vous-même ? Pouvez-vous collaborer avec d'autres acteurs ?
3. Analyse institutionnelle	Dans quelle mesure votre institution est-elle prête au changement en termes de contenu et de processus ?
4. Analyse des parties prenantes et sensibilisation	Qui travaillera et utilisera les objets numériques ? Qui décide du financement ? Engagez-vous avec ces parties prenantes en utilisant un langage et des termes qu'elles comprendront.
5. Objectifs : buts du champ d'application, activités, plan et coûts	Déterminez ce que vous allez faire, qui fera quoi, combien cela coûtera et quand cela se produira.
6. Définir les avantages pour la stratégie organisationnelle	Assurez-vous d'exprimer les avantages de votre étude de cas de manière à ce que vos financeurs la comprennent.
7. Quels sont les autres besoins ?	Faut-il inclure une analyse coûts-avantages ou une liste d'options basée sur les dépenses et les résultats ?
8. Valider et affiner l'étude de cas	Examinez et testez votre étude de cas par rapport aux meilleures pratiques et identifiez les autres besoins.
9. Réaliser l'étude de cas avec un maximum d'impact	Avez-vous un expert reconnu à solliciter au sein de l'organisation ? N'oubliez pas que vous pourrez avoir besoin de le mobiliser de nouveau.
10. Partager une étude de cas achevée	Retirez les données confidentiels et partagez l'étude de cas en ligne afin que d'autres personnes puissent bénéficier de votre travail.

Pour un modèle générique d'étude de cas sur la préservation numérique et pour plus d'informations, voir le *Digital Preservation Business Case Toolkit*[30].

Avantages

Les avantages sont associés aux coûts et aussi aux risques (voir Gestion des risques et des changements). Si les risques sont atténués, ils deviennent une sorte d'avantage. L'objectif de l'acquisition de tout objet numérique repose dans son utilisation. Les usages pour lesquels l'objet numérique est mis à disposition représentent un avantage pour ces utilisateurs. Si une organisation doit comprendre les coûts associés aux objets numériques, elle doit également comprendre les avantages. Les avantages peuvent être utilisés pour justifier les coûts par l'élaboration de plan d'activités.

La mesure des avantages est souvent assez difficile, surtout lorsque ces avantages ne se prêtent pas facilement à une expression quantitative. Souvent, il faudra recourir à plusieurs approches pour analyser les résultats qualitatifs et quantitatifs et présenter les différences constatées. Pour aider les institutions, le projet *Keeping Research Data Safe* (KRDS) a créé le *KRDS Benefits Framework*[31] et le *KRDS Benefits Analysis Toolkit*[32]. Ces outils visent à aider les institutions à identifier l'ensemble des avantages découlant de la gestion et de la préservation des données de recherche, et à les présenter de manière succincte à des parties prenantes différentes (par exemple, lors de l'élaboration d'études de cas ou de la sensibilisation). La boîte à outils est également facilement applicable aux avantages de la préservation numérique orientée vers d'autres catégories d'objets numériques.



Le *KRDS Benefits Framework* utilise trois dimensions pour mettre en lumière les avantages que les investissements peuvent générer. Ces dimensions servent de cadre de haut niveau dans lequel la réflexion sur les bénéfices peut être organisée, puis ensuite affinée en propositions de valeur plus ciblées à l'aide de la boîte à outils. Ce processus vous aide à identifier les changements que vous essayez d'apporter, les résultats, les bénéficiaires et le temps nécessaire pour réaliser ces avantages.

[30] Kit d'étude de cas sur la préservation numérique
[31] Cadre des avantages du projet KRDS
[32] Boîte à outil pour l'analyse des avantages

Coûts

Une étude de cas ne portera normalement pas uniquement sur le coût de mise en place de la solution de préservation numérique, mais aussi sur le coût global, y compris les coûts de gestion du projet/programme et les autres activités entreprises pour soutenir la mise en œuvre, telles que la formation et la publicité. Cependant, les coûts de la préservation numérique constituent souvent l'élément le plus critique.

Pourquoi comprendre les coûts de la préservation numérique ?

Voici quelques raisons pour lesquelles une organisation pourrait vouloir estimer les coûts de la préservation numérique :

- planifier et budgétiser la création d'un nouvel entrepôt numérique ou l'extension d'un entrepôt existant ;
- ajouter un nouvel objet numérique à votre entrepôt numérique et décider si cette solution est rentable maintenant ou à long terme ;
- fournir un point de comparaison avec des organisations similaires afin d' adopter les gains d'efficacité qui ont été identifiées par d'autres ;
- décider d'externaliser les activités ou de les conserver en interne ;
- décider du montant à facturer pour la fourniture d'un service de préservation numérique aux clients ;
- comprendre où les ressources sont utilisées ou sous-utilisées et les domaines où une allocation supplémentaire de ressources pourrait être bénéfique.

Qu'est-ce que la modélisation du coût du cycle de vie ?

Un certain nombre de projets de recherche et de développement ont cherché à modéliser les coûts de la préservation numérique tout au long du cycle de vie, de la création et de l'ingestion à la préservation et, finalement, à l'accès. Pour un grand nombre de projets, la compréhension de ce travail, la recherche des résultats les plus applicables à une situation particulière, le choix d'un modèle et sa mise en pratique représentent un défi important. Le projet 4C a étudié, analysé et évalué ce travail, et fournit des conseils pour en tirer le meilleur parti :

- La partie « *Starting out with curation costs* » (Commencer par les coûts de préservation) fournit une introduction aux concepts.
- La partie « *Using cost models* » (Utiliser des modèles de coûts) décrit comment choisir un modèle de coûts approprié à votre organisation.
- La partie « *Cost Concept Model and Gateway Specification* » (Modèle de concept de coût et spécification de la passerelle) fournit plus de détails, y compris un guide pour développer un modèle selon vos propres exigences en examinant des concepts tels que le « risque », la « valeur », la « qualité » et la « pérennité ».

Défis de la modélisation des coûts

La modélisation des coûts a été identifiée comme une activité particulièrement difficile, avec un certain nombre d'aspects complexes, tels que :

- L'articulation des facteurs ou des objectifs pour le calcul des coûts de la préservation numérique.
- La préservation numérique est une cible mouvante, définie par l'évolution des technologies et des exigences institutionnelles.

- Le niveau de détail : à un niveau élevé, la modélisation devient moins utile car elle porte généralement sur un échantillon de différents contextes de préservation. À un niveau inférieur, la modélisation devient rapidement très complexe, ce qui rend les modèles difficiles à développer, à maintenir et à mettre en œuvre.
- Les organisations sont réticentes à partager les données de calcul des coûts, avec lesquelles les modèles peuvent être élaborés et validés.
- Même lorsque les données de calcul des coûts ont été partagées, il est souvent difficile d'établir une correspondance entre les différents modèles de calcul des coûts auxquels elles sont associées.
- Il est difficile, et parfois inutile, de séparer les coûts de la préservation numérique des autres coûts d'activité (par exemple, la numérisation des collections de manière à faciliter l'intégration du contenu dans l'environnement de préservation : s'agit-il de coûts de préservation numérique ou de numérisation ?).

C'est pourquoi la modélisation des coûts de la préservation numérique sur l'ensemble du cycle de vie est une activité qu'il convient d'aborder avec prudence. La modélisation des coûts sera toujours une approximation et vous devez donc décider du temps que vous êtes prêt à y consacrer pour obtenir une réponse moins approximative.

Gestion des coûts

Il est possible de gérer les coûts grâce à une planification minutieuse. L'un des moyens consiste à bien concevoir les processus. Les modes de création ou d'acquisition, de gestion et de diffusion des objets numériques entraînent des coûts. Ces coûts sont à la discrétion de l'organisation et peuvent être gérés. Le processus de bout en bout, de l'acquisition à la diffusion, doit être conçu pour que toutes les activités soient aussi efficaces que possible. Toutes les étapes doivent être conçues de manière à minimiser le besoin de ressources, tout en maximisant l'efficacité. Si l'efficacité fonctionne bien à grande échelle, un processus efficace ne doit pas nécessairement être un processus à grand volume. L'automatisation des étapes systématiques peut également permettre de gagner du temps et de mettre en place des processus efficaces et cohérents. Les coûts initiaux de conception et de mise en œuvre des processus peuvent être compensés par une rentabilité à plus long terme.

Impact

L'impact signifie généralement la mesure des avantages, en particulier pour le grand public et la société, qui est effectuée après la réalisation d'un projet d'étude de cas.

Pour les projets et les études de cas de petite taille, l'impact peut être un simple ensemble de mesures, telles que les téléchargements ou le nombre de requêtes enregistrées sur un site web par rapport auquel la réussite peut être facilement évaluée.

Pour les projets et programmes plus importants, il peut s'agir d'une évaluation plus approfondie visant à justifier les ressources dépensées. Elle peut comprendre un mélange de mesures quantitatives et qualitatives, et sera normalement réalisée par des spécialistes externes travaillant avec le personnel en charge de l'entrepôt numérique. Ils utilisent des méthodes issues de l'économie, de la gestion et des sciences de l'information, par exemple l'analyse coûts-avantages ou l'évaluation contingente, ainsi que des méthodes traditionnelles des sciences sociales, telles que les entretiens, les enquêtes et les groupes de discussion.

L'évaluation de l'impact implique le choix de mesures ou d'indicateurs et nécessite une planification minutieuse et un accord sur ce qui doit être mesuré et comment. Les mesures font souvent appel à des éléments facilement dénombrables, tels que les téléchargements, ou à des mesures d'échelle qui ne sont pas réellement numériques, comme les échelles de notation ou les catégories de variables. Il y a généralement un compromis entre ce qui devrait idéalement être mesuré (par exemple, les utilisateurs et l'utilisation) et les mesures indirectes qui sont faciles à saisir et à mesurer (par exemple, les « visiteurs uniques » et les téléchargements sur le web).

Ressources

Sustainable Economics for a Digital Planet: Ensuring Long-Term Access to Digital Information
http://brtf.sdsc.edu/biblio/BRTF_Final_Report.pdf
Le groupe d'experts de la *Blue Ribbon Task Force* a étudié la préservation et l'accès numériques durables d'un point de vue économique. Ce rapport final identifie les problèmes intrinsèques à tous les objets numériques pérennisés et propose des actions que les parties prenantes peuvent prendre pour relever ces défis de la pérennité. Des programmes d'action ciblent les principaux groupes de parties prenantes et des stratégies de préservation spécifiques à chaque domaine (2010, 116 pages).

The Value and Impact of Data Sharing and Curation: A synthesis of three recent studies of UK research data centres
http://repository.jisc.ac.uk/5568/1/iDF308_-_Digital_Infrastructure_Directions_Report%2C_Jan14_v1-04.pdf
Cette synthèse résume et reflète les résultats combinés d'une série d'enquêtes indépendantes sur la valeur et l'impact de trois centres ou services de données de recherche bien établis au Royaume-Uni (*Economic and Social Data Service*, *Archaelogy Data Service* et *British Atmospheric Data Centre*). Les études ont adopté un certain nombre d'approches pour explorer la valeur et l'impact des services de données de recherche, ainsi que le partage et l'archivage des données que ces services ont facilités. La collecte de données a rassemblé des enquêtes ciblées auprès des utilisateurs et des déposants, ainsi que des données financières et opérationnelles des centres de données (par exemple, enregistrements des utilisateurs, dépôts et téléchargements d'ensembles de données), complétées par des entretiens approfondis. Tous les impacts ne peuvent pas être saisis et quantifiés ; c'est pourquoi ces approches économiques ont été utilisées en parallèle avec d'autres outils, comme le *KRDS Benefits Framework*, pour illustrer des avantages plus larges (2014, 26 pages).

Jisc Guide: Digitising your collections sustainably

https://www.jisc.ac.uk/guides/digitising-your-collections-sustainably[33]

Le guide fournit un point de départ pour examiner les questions nécessaires pour créer et construire un modèle de gestion qui soutiendra la pérennité de la numérisation et des collections numériques.

4C Project Collaboration to Clarify the Costs of Curation

https://www.4cproject.eu/

Le projet 4C, financé par l'Union européenne, visait à aider les organisations de toute l'Europe à investir plus efficacement dans la préservation numérique. Une série de rapports et de ressources ont été produits et sont disponibles sur la page Outputs and Deliverables. Il s'agit notamment du modèle de pérennité de la préservation numérique (*Digital Curation Sustainability Model)*, d'une évaluation des modèles de coûts et d'une analyse des besoins et des lacunes (*Evaluation of Cost Models and Needs and Gaps Analysis)*, d'un rapport sur les risques, les avantages, l'impact et la valeur *(Report on Risk, Benefit, Impact and Value)*, et d'un projet de modèle de référence de pérennité économique *(Draft Economic Sustainability Reference Model)*. Le rapport sur l'évaluation des modèles de coûts évalue dix modèles de coûts disponibles, dont le *KRDS* et *LIFE*. Représentant une autre contribution importante, *The Curation Costs Exchange* (*CCEx*)[34] est une plateforme communautaire qui aide les organisations de toute sorte à évaluer les coûts des pratiques de préservation par la comparaison et l'analyse. Le *CCEx* vise à fournir des informations réelles sur les coûts afin d'aider à faire des investissements mieux informés dans la préservation numérique. Le *CCEx* a été lancé en 2014 par le projet 4C et est maintenant maintenu et régi par la *Digital Preservation Coalition* avec l'aide de nestor, le centre allemand de préservation numérique, et de la *Netherlands Coalition for Digital Preservation* (*NCDD*).

Digital Preservation Business Case Toolkit

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

Cette boîte à outils fournit un guide approfondi pour la rédaction d'une étude de cas axée sur les activités de préservation numérique. Cette ressource s'adresse aux professionnels (et à leurs cadres) qui travaillent avec des ressources numériques et qui souhaiteraient obtenir des fonds pour développer leurs activités de préservation numérique. La boîte à outils est principalement destinée à ceux qui cherchent à obtenir des fonds supplémentaires au sein de leur organisation, mais elle peut également fournir des informations utiles à ceux qui rédigent une demande de fonds pour un projet auprès d'un organisme de financement externe. Cette ressource comprend un guide détaillé pour élaborer une étude de cas, ainsi qu'un modèle pour la réalisation d'une étude de cas. Créé par le projet SPRUCE financé par le Jisc en 2013, le wiki de la boîte à outils est hébergé par la DPC.

Keeping Research Data Safe (KRDS) Benefits Toolkit

http://www.beagrie.com/krds/

Keeping Research Data Safe (KRDS) est une série d'études, d'outils et de méthodologies qui se concentrent sur les défis de l'évaluation des coûts et des avantages de la préservation des données de recherche. Bien que centrés sur les données de recherche, les outils sont facilement adaptables pour s'appliquer à d'autres domaines de la préservation numérique. Les résultats disponibles

[33] NdT : Le guide n'est plus disponible sur le site du Jisc à la date de la présente traduction. Le lien est maintenu pour prendre en compte une éventuelle mise à jour par le Jisc.

[34] Échange des coûts de préservation

comprennent une fiche d'information sur le KRDS (*KRDS Factsheet)*, un guide de l'utilisateur du KRDS *(KRDS User Guide)*, un modèle de coût d'activité du KRDS *(KRDS Activity Cost Model)* et une boîte à outils d'analyse des avantages du KRDS *(KRDS Benefits Analysis Toolkit)*, ainsi que des ressources et des rapports supplémentaires. Les projets KRDS entre 2008 et 2011 ont été financés par Jisc.

20 Cost Questions for Digital Preservation

http://www.metaarchive.org/public/publishing/ma_20costquestions_final.pdf?thumblink

L'organisation *MetaArchive Cooperative* a produit une série de 20 questions pour « aider les institutions dans leur analyse comparative des différentes solutions de préservation numérique ». Ce travail marque une évolution de l'élaboration de modèles de coûts prédictifs détaillés vers une approche plus générale qui cherche à identifier et à comprendre les principaux facteurs de coûts, plutôt que les coûts réels eux-mêmes.

Blog DSHR

http://blog.dshr.org/search/label/storage%20costs
http://blog.dshr.org/search/label/cloud%20economics

David Rosenthal est un blogueur assidu sur le sujet des coûts de stockage, considérant souvent l'impact de l'évolution des technologies de stockage sur les coûts de préservation et sur le stockage dans le *cloud*.

A Digital Asset Sustainability and Preservation Cost Bibliography

http://blogs.loc.gov/digitalpreservation/2012/06/a-digital-asset-sustainability-and-preservation-cost-bibliography/

Cette bibliographie « rassemble des articles sur l'évaluation contingente, l'évaluation des écosystèmes et les coûts généraux de la connaissance, ainsi que des articles qui traitent directement des questions de coût associées à la préservation et à la gérance numériques ».

Digital Preservation and Data Curation Costing and Cost Modelling

http://wiki.opf-labs.org/display/CDP/Home

Liste de modèles de coûts de préservation numérique et d'initiatives de modélisation des coûts.

The Cost of Inaction Calculator Rationale

https://coi.weareavp.com/rationale

Il s'agit d'une excellente vidéo d'information d'*AVPreserv* sur le coût de l'inaction et la justification de rentabilité en matière de préservation numérique. Elle est centrée sur les objets audiovisuels, mais il vaut la peine d'écouter et de réfléchir latéralement à la justification sous-jacente, quel que soit le type d'objet numérique que vous détenez (8 mins 41 sec).

Études de cas

Études de cas provenant de la boîte à outils KRDS Benefits Toolkit

Il existe quatre études de cas fournissant des exemples concrets de fiches de travail remplies par les partenaires du projet, comme suit :

Archéologie

http://www.ukoln.ac.uk/events/i2s2-krds/presentations/catherine-hardiman-krds-benefit-framework-2011-07-v2.ppt

Le contexte de cette étude de cas est fourni dans la présentation de l'atelier de diffusion de l'*Archaeology Data Service* (ADS). Des exemples de réalisation sont disponibles dans la feuille de travail sur le cadre des avantages de l'ADS, *ADS Benefits Framework Worksheet* (PDF), et dans la feuille de travail sur la chaîne de valeur et l'impact de l'ADS, *ADS Value-chain and Impact Worksheet* (Excel 97-2003).

Santé : Études de cohorte de la population

http://www.ukoln.ac.uk/events/i2s2-krds/presentations/dipak-kalra-krds-benefits-2011-07.ppt

Le contexte de cette étude de cas est fourni dans la présentation de l'atelier de diffusion des études de cohorte du *Medical Research Council*. Un exemple de réalisation est disponible dans la feuille de travail sur la chaîne de valeur et l'impact des études de cohortes, *Cohort Studiers Value-chain and Impact Worksheet* (Excel 97-2003).

Citation des données de recherche : SageCite

http://www.ukoln.ac.uk/events/i2s2-krds/presentations/monica-duke-krds-sagecite-benefits-2011-07.ppt

Le contexte de cette étude de cas est fourni dans la présentation de l'atelier de diffusion de *SageCite*. Un exemple de réalisation est disponible dans la feuille de travail du cadre des avantages de *SageCite*, *SageCite Benefits Framework Worksheet* (PDF).

Sciences sociales : *UK Data Archive* (UKDA)

http://www.ukoln.ac.uk/events/i2s2-krds/presentations/matthew-woollard-krds-benefits-2011-07.ppt

Le contexte de cette étude de cas est fourni dans la présentation de l'atelier de diffusion de l'*UKDA*. Un exemple de réalisation est disponible dans la feuille de travail sur l'impact des avantages de l'*UKDA, UKDA Benefits Impact Worksheet* (PDF).

Digital Preservation Business Case Toolkit

Les quatre études de cas suivantes proviennent d'activités menées dans le cadre des prix du projet *SPRUCE* :

Étude de cas de la bibliothèque *Bishopsgate*

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

Audit des collections et étude de cas axés sur les premières étapes de la préservation numérique.

Étude de cas : *Institute of education*

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

Examen de l'approche et de l'élaboration d'une étude de cas pour l'archivage des documents administratifs numériques.

Étude de cas : *Northumberland estates*

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

Évaluation des solutions d'entrepôt numérique et étude de cas associée pour la préservation numérique.

Étude de cas : *Lovebytes*
http://wiki.dpconline.org/index.php?title=Lovebytes_case_study
Essai de stabilisation des médias et de préservation du contenu, et étude de cas pour passer à un état de production.

Références

Keeping Research Data Safe (KRDS), 2011. *Digital Preservation Benefits Analysis Toolkit*. Disponible à l'adresse suivante : http://beagrie.com/krds-i2s2.php

Activités organisationnelles

Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

A qui s'adresse-t-il ?

Les créateurs et les éditeurs de ressources numériques, les prestataires de services tiers, les encadrants (*DigCurV : Manager Lens*) et le personnel opérationnel (*DigCurV : Practitioner Lens*) chargés de mettre en œuvre des activités institutionnelles pertinentes pour la préservation numérique. Il est supposé que celles-ci comprendront a) du personnel provenant de parties structurellement séparées de l'organisation, et b) un large éventail de connaissances en matière de préservation numérique, du niveau débutant au niveau expert ; c) des perspectives techniques et non techniques ; d) un large éventail d'activités fonctionnelles ayant un lien direct ou indirect avec les activités de préservation numérique.

Niveau de connaissance présumé

L'éventail est large, du novice au confirmé.

Objectif

- Fournir des indications sur les sources de conseils et d'orientation visant à encourager les bonnes pratiques en matière de création et de gestion des objets numériques. L'importance du producteur pour faciliter la préservation numérique est soulignée tout au long du manuel mais particulièrement dans la section Création d'objets numériques. Les bonnes pratiques en matière de numérisation et de création d'autres objets numériques sont cruciales pour la viabilité continue des objets numériques.
- Sensibiliser aux facteurs qui doivent être pris en compte lors de la création ou de l'acquisition d'objets numériques.
- Fournir des indications sur les sources de conseils et d'orientation utiles, tant pour les novices que pour ceux qui ont déjà commencé à réfléchir aux implications de la technologie numérique sur leurs activités opérationnelles.

Création d'objets numériques

Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Introduction

« La première ligne de défense contre la perte d'informations numériques de valeur repose sur les producteurs, les fournisseurs et les propriétaires d'informations numériques ». (Waters et Garrett, 1996)

Le groupe de travail sur l'archivage de l'information numérique est le premier à avoir souligné l'importance du rôle du producteur pour aider à garantir l'accès à long terme aux ressources numériques qu'il crée. Ce point de vue a été réitéré dans de nombreux autres documents depuis lors.

Il est clair qu'il ne peut pas être attendu de la plupart des producteurs individuels qu'ils s'engagent à long terme à pérenniser le contenu numérique qu'ils créent au-delà de leurs besoins de gestion. Chaque ressource numérique a un cycle de vie et différents acteurs et intérêts au sein de celui-ci. Cependant, il est à la fois hautement souhaitable et réalisable qu'un dialogue s'établisse entre les dépositaires à long terme et les producteurs lorsque des questions de préservation à long terme sont en jeu. Il est souvent aussi dans l'intérêt du producteur que le contenu créé soit bien formé, complet, correct et utilisable à des fins actuelles et futures. Étant donné le rôle crucial du producteur pour réaliser la préservation à court et moyen terme, souvent pendant plusieurs décennies, et au moins pour faciliter la préservation à moyen et long terme, il est essentiel d'encourager les bonnes pratiques (et de sensibiliser les dépositaires).

Cette section est uniquement dédiée à encourager les bonnes pratiques dans la création d'objets numériques pour contribuer à leur longévité d'utilisation active, de gestion dans le futur et de préservation. Vous devez vous référer à d'autres sections pertinentes du manuel pour les activités et les conseils qui s'y rapportent.

Nous nous concentrons sur les implications génériques pour la préservation numérique au moment du processus de numérisation (création de substituts numériques) ou celui de création d'objets numériques natifs.

La création de substituts numériques ou de fichiers nativement numériques spécifiques à un domaine, tels que les documents d'activité électroniques, les données de recherche ou les informations numériques personnelles, fait l'objet d'excellentes sources de conseils et d'orientation supplémentaires. Les principales références sont fournies dans la section Ressources et études de cas.

Créer des objets numériques natifs

La préservation numérique fait référence à la série d'activités jugées nécessaires pour garantir un accès continu aux objets numériques aussi longtemps que nécessaire. Cela comprend les activités

de création d'objets numériques natifs nécessaires pour répondre aux besoins permanents du producteur original.
Souvent, nombre de ces actions nécessaires pour assurer un accès continu à long terme chevauchent les meilleures pratiques adaptées aux besoins immédiats de gestion. En effet, de nombreuses organisations et personnes créent des objets numériques qu'elles auront besoin d'utiliser et gérer pendant de nombreuses décennies. Elles ne se considèrent probablement pas comme des acteurs de la préservation numérique et d'autres termes tels que « continuité numérique » sont fréquemment utilisés pour communiquer sur la manière dont ces actions les affectent lorsqu'ils ne sont pas des organismes de mémoire tels que des musées, des bibliothèques ou des services d'archives ayant une mission de préservation.
Il est important que les producteurs réalisent que s'ils ne travaillent pas activement pour assurer la continuité, leurs objets numériques peuvent facilement devenir inutilisables. Il s'agit de s'assurer que leurs informations sont complètes, disponibles et donc utilisables pour leurs besoins de gestion.
Vos informations sont utilisables si vous pouvez :

- les trouver quand vous en avez besoin ;
- les ouvrir quand vous en avez besoin ;
- les utiliser de la manière dont vous en avez besoin ;
- comprendre ce qu'elles sont et de quoi il s'agit ;
- avoir confiance dans le fait qu'elles sont ce qu'elles disent être.

Cela vous permet d'agir de manière responsable, légale, efficace et efficiente et vous aide à protéger votre réputation, à prendre des décisions éclairées, à réduire les coûts et à améliorer votre offre de service. Pour plus d'informations sur les premières étapes de la préservation numérique, voir la section Pour commencer.

Le tableau suivant fournit des conseils sur les questions et les actions clés à prendre en compte lors de la création d'objets numériques afin de garantir leur longévité d'utilisation active et leur préservation potentielle à long terme.

Pérenniser des objets numériques natifs	
1	Logiciels et formats
	Choisissez un logiciel qui est bien pris en charge et qui crée des fichiers pouvant être lus par une variété de programmes différents. Voir la section Formats de fichiers et normes du manuel pour des conseils pertinents.
2	Noms de fichiers
	Utilisez un nom de fichier descriptif court du contenu et de la date qui fournit un contexte et peut être facilement compris par les humains et les ordinateurs, maintenant et à l'avenir. N'utilisez pas d'espaces ou de caractères spéciaux (autres que – ou _), cela évitera toute mauvaise interprétation par le matériel ou les logiciels informatiques. Mettez les informations de date dans le format normalisé par l'ISO 8601:2004 : YYYY-MM-DD. Cela permet de disposer d'une méthode cohérente pour le suivi des versions. Notez que les métadonnées de date de fichier séparées générées par les systèmes peuvent souvent changer automatiquement lors d'actions ultérieures. Utilisez une méthode cohérente pour afficher les versions des fichiers. Il peut s'agir de la date comme ci-dessus, complétée si nécessaire, par exemple par un numéro de version v1, v2, v_final, etc.
3	Stockage et sauvegarde
	Voir les principes d'utilisation des systèmes de stockage informatique pour la préservation

	numérique dans la section Stockage du manuel pour des conseils pertinents.
4	Prenez connaissance de vos obligations et des pratiques les plus pertinentes.
	Voir la section Respect de la législation du manuel pour des conseils pertinents. De nombreuses obligations et meilleures pratiques seront spécifiques à un projet ou à un secteur – voir l'encadré « Créer des données de recherche » ci-dessous pour un exemple.
5	Planifier les transitions
	Certaines transitions peuvent être prévues et planifiées, d'autres peuvent être imprévues mais peuvent être atténuées par une bonne planification et de bonnes procédures. Voir les ressources et les études de cas ci-dessous ainsi que les sections Planification de la préservation et Gestion des risques et du changement du manuel pour des conseils pertinents.

Créer des substituts numériques

L'accent mis sur la numérisation dans cette section reflète son importance actuelle alors qu'un nombre croissant d'institutions se lancent dans la numérisation d'une partie de leurs collections. Il est important de souligner que ce manuel n'envisage pas le potentiel de la numérisation comme outil de reformatage de la préservation. L'accent est mis sur la préservation des objets numériques natifs ou des produits de la numérisation (les substituts numériques eux-mêmes), et non sur la préservation des originaux analogiques.

Une exception importante est celle des objets audiovisuels. Les objets audio et vidéo ont besoin de la numérisation pour la survie même de leur contenu, en raison de l'obsolescence des équipements de lecture et de la détérioration et de l'endommagement des éléments physiques, qu'ils soient analogiques ou numériques (voir Images animées et son).

De nombreux projets de numérisation citent l'accès amélioré comme objectif principal, un objectif parfaitement légitime, mais si l'on n'accorde pas le soin et l'attention nécessaires à la manière dont cet accès peut être maintenu dans le temps, il ne sera pas pérenne. Il est peu probable que l'ensemble des initiatives de numérisation actuelles tiennent compte de la viabilité à long terme des substituts numériques qu'elles créent. Il est donc utile d'encourager les bonnes pratiques en matière de création d'objets numériques et d'indiquer les sources d'orientation existantes.
Une bonne partie de ce qui est aujourd'hui numérisé a commencé sa vie en tant que contenu numérique natif. Il a été converti en un format analogique tel que l'impression sur papier avant que la nécessité d'un accès numérique et d'une re-numérisation ne soit reconnue. Ce cycle doit rapidement évoluer vers une simple gestion d'un plus grand nombre de contenus numériques natifs.

Pérenniser les substituts numériques : considérations sur la préservation numérique	
1	Évaluation des besoins de numérisation
	L'objet a-t-il déjà été numérisé ? Si oui, est-il conforme à une norme appropriée et facilement accessible à votre public ?
2	Trouver des fonds pour le projet
	Quelles sont les politiques d'archivage existantes, tant de la part de l'organisme de financement (s'il s'agit d'un financement externe) que de l'institution ayant la responsabilité première du projet ?
3	Planifier un projet et affecter des ressources
	Il est nécessaire de prévoir des fonds récurrents pour la maintenance des copies numériques ainsi que des fonds ponctuels pour la conversion. S'assurer que toutes les parties prenantes concernées sont au courant du projet (par

	exemple, si une autre partie de l'organisation ou une agence externe est censée assurer la maintenance de la ressource, elles devront être incluses dans les discussions à ce stade, si ce n'est avant). Identifier une stratégie pour faire avancer les acquis du projet de manière durable après que le projet ait atteint ses objectifs. Cette stratégie pourrait consister à intégrer les produits du projet dans le catalogue des collections de l'organisation mère, ou à désigner une institution partenaire pour recevoir ces produits.
4	Sélection des objets
	Droits d'auteur. Il sera nécessaire de s'assurer que l'autorisation est donnée à la fois pour numériser l'original et pour faire des copies de la copie numérique à des fins de préservation et de communication. Pour plus d'informations, voir Conformité juridique. État et complétude de l'original. Est-il possible de le scanner à nouveau ultérieurement si la copie numérique est perdue ?
5	Décider de la manière dont le contenu de l'information doit être organisé (par exemple, bases de données de textes consultables et/ou images de pages de documents)
	Sélection des formats de fichier et de stockage appropriés pour le master, les copies d'archives et les dérivés, voir Formats de fichier et normes, Métadonnées et documentation et Stockage.
6	Décider de la méthode de numérisation appropriée à l'original analogique et aux objectifs du projet.
	Préparer les originaux pour la numérisation. Les détails de la méthode de numérisation doivent être documentés et joints au document de métadonnées pour permettre une gestion future.
7	Préparer les originaux pour la numérisation
	Les Archives nationales du Royaume-Uni fournissent des normes et des conseils sur la préparation des documents en vue de leur numérisation (Archives nationales du Royaume-Uni, 2015). Les originaux seront-ils conservés ? Ne prenez aucune mesure pour vous débarrasser des originaux tant qu'il n'est pas établi que : a) la version électronique est légalement admissible et/ou b) la version électronique est capable d'être préservée à long terme. La décision de conserver ou non les originaux après la numérisation ne sera bien sûr pas un problème pour les projets de numérisation de trésors précieux au sein d'une collection, la question principale sera alors de savoir si l'original est trop fragile pour être numérisé à une date ultérieure si la copie numérique est perdue. Dans tous ces cas, si la copie numérique devient le principal moyen d'accès, elle sera soumise aux mêmes exigences que l'objet numérique natif.
8	Conversion
	Documentation des caractéristiques techniques. Algorithme de compression (si utilisé) ; profondeur de bits requise ; résolution de balayage, etc. Créez des copies de sauvegarde dès que la conversion est entreprise.
9	Contrôles d'assurance qualité
	Le substitut numérique doit être d'une qualité de préservation acceptable. Si l'on fait appel à des services tiers, il faut veiller à ce que la documentation clarifie la responsabilité en matière d'assurance qualité.
10	Indexation et catalogage final
	Métadonnées pour la recherche des ressources et pour la gestion et la préservation de la copie numérique.
11	Charger des données dans les systèmes informatiques Exigences en matière de stockage des documents pour les copies d'accès et de préservation

	(si elles sont différentes). Faites des copies de sauvegarde, le cas échéant.
12	Mettre en œuvre des stratégies d'archivage et de préservation ou un transfert à un organisme en charge de la préservation
	Normes requises pour les formats, les supports de stockage, la documentation et les procédures de transfert. Stockage des originaux et des copies de sauvegarde. Stratégies pour le rafraîchissement des supports et les changements dans l'environnement technologique.

Ressources

Digitisation at The National Archives
http://nationalarchives.gov.uk/documents/information-management/digitisation-at-the-national-archives.pdf
Ce document définit les normes et les exigences des Archives nationales du Royaume-Uni pour la numérisation des documents analogiques de sa collection. Il est aussi recommandé aux ministères du gouvernement du Royaume-Uni qui souhaite numériser n'importe quel type de documents papier. Il couvre : l'ensemble du processus de numérisation, depuis le balayage initial jusqu'à la livraison des images pour la préservation, y compris les spécifications des images scannées des Archives nationales ; le balayage des documents dont les images résultantes deviendront le document public légal pour une préservation permanente ; et le balayage des documents dont les images résultantes deviendront des substituts numériques, les documents papier originaux étant conservés et restant le document public légal (juillet 2015, 56 pages).

Koninklijke Bibliotheek[35] : *Metamorfoze preservation imaging guidelines*
https://www.metamorfoze.nl/sites/default/files/publicatie_documenten/Metamorfoze_Preservation_Imaging_Guidelines_1.0.pdf
Metamorfoze est le programme national des Pays-Bas pour la préservation du patrimoine du papier. Les directives sont destinées à la numérisation des objets bidimensionnels tels que les manuscrits, les archives, les livres, les journaux et les magazines. Elles peuvent également être appliquées aux photographies, peintures et dessins techniques. Les lignes directrices portent exclusivement sur la qualité de l'image et les métadonnées du fichier master de préservation, à partir duquel tous les résultats destinés à l'impression et/ou au web peuvent être dérivés (2012, 44 pages).

Preparing Collections for Digitization
Ce livre de 2010, écrit par Anna E. Bulow et Jess Ahmon, offre des conseils pratiques sur le processus de numérisation des collections de bout en bout et peut être utilisé comme un manuel de référence pratique pour les gestionnaires de collections qui se lancent dans un projet de numérisation ou qui gèrent un projet existant. Il couvre également certaines questions plus larges telles que l'utilisation de substituts pour la préservation et la durabilité à long terme de l'accès numérique (208 pages).

[35] Bibliothèque nationale des Pays-Bas.

InterPARES 2 Creator Guidelines: Making and Maintaining Digital Materials
http://www.interpares.org/ip2/display_file.cfm?doc=ip2(pub)creator_guidelines_booklet.pdf
Cette brochure fournit des conseils aux personnes qui créent des objets numériques dans le cadre de leurs activités professionnelles et personnelles afin de les aider à assurer leur préservation (10 pages).

InterPARES 2 Preserver Guidelines: Preserving Digital Records: Guidelines for Organisations
http://www.interpares.org/public_documents/ip2(pub)preserver_guidelines_booklet.pdf
Cette brochure fournit des conseils à toute organisation responsable de la préservation à long terme de documents numériques (10 pages).

Ressources sur les supports numériques de l'organisation Jisc
Intellectual property rights in a digital world
https://www.jisc.ac.uk/guides/intellectual-property-rights-in-a-digital-world

Digitising your collections sustainably
https://www.jisc.ac.uk/guides/digitising-your-collections-sustainably[36]

Federal Agencies Digitization Guidelines Initiative
http://www.digitizationguidelines.gov/
Il s'agit d'un effort de collaboration entre les agences fédérales américaines pour définir des directives, des méthodes et des pratiques communes pour la numérisation des contenus historiques. Dans ce cadre, deux groupes de travail étudient les questions spécifiques à deux domaines majeurs, l'image fixe et l'audiovisuel.

Future Proof – Protecting our digital future
http://futureproof.records.nsw.gov.au/
Future Proof est une initiative des State Records du gouvernement de l'État de Nouvelles-Galles du Sud en Australie. Ce site web et ce blog couvrent les produits et les projets des State Records qui concernent spécifiquement les documents numériques. Les liens des catégories offrent des passerelles utiles vers différentes ressources hébergées sur le site.

The Curation Reference Manual
http://www.dcc.ac.uk/resources/curation-reference-manual
Cette ressource gérée par le *Digital Curation Centre* contient des conseils, des informations approfondies et des critiques sur les techniques de préservation numérique actuelles et les meilleures pratiques. Le manuel est un projet continu, mené par la communauté, qui implique des membres de la communauté du DCC qui suggèrent des sujets, créent des parties du manuel et organise des évaluations par des pairs. Chaque partie du manuel est conçue pour aider les dépositaires, les producteurs et les utilisateurs de données à mieux comprendre les défis auxquels ils sont confrontés et le rôle qu'ils jouent dans la création, la gestion et la préservation des informations numériques au fil du temps.

[36] NdT : Le guide n'est plus disponible sur le site du Jisc à la date de la présente traduction. Le lien est maintenu pour prendre en compte une éventuelle mise à jour par le Jisc.

UK Data Archive : Consentement au partage des données
https://www.ukdataservice.ac.uk/manage-data/legal-ethical/consent-data-sharing
La collecte, l'utilisation et le partage de données dans le cadre de la recherche avec des personnes exigent le respect d'obligations éthiques et juridiques. Des lois telles que la loi sur la protection des données[37], la loi sur la liberté de l'information38 et la loi sur les statistiques et les services d'enregistrement[39]régissent également l'utilisation de certains types de données. Le présent guide offre une aide sur la manière dont les données de recherche peuvent être partagées tout en assumant ses responsabilités éthiques ou juridiques.

An Elevator Pitch for File Naming Conventions
http://acrl.ala.org/techconnect/post/an-elevator-pitch-for-file-naming-conventions
Cet article du blog TechConnect de l'*Association of College and Research Libraries* (ACRL) plaide en faveur de l'adoption d'une approche cohérente lors du nommage de fichiers numériques ou de composants logiciels, en montrant les conséquences de ne pas y procéder. (2013).

Guide de la continuité numérique des Archives nationales du Royaume-Uni
http://www.nationalarchives.gov.uk/information-management/manage-information/policy-process/digital-continuity/
Des conseils détaillés sur la continuité numérique des Archives nationales du Royaume-Uni sont particulièrement utiles les publications suivantes : *Understanding Digital Continuity*, une introduction au sujet (2011, 20 pages), et *Managing Digital Continuity*, qui vous emmène à travers un processus en quatre étapes : 1) Planifier les actions 2) Définir vos besoins de continuité numérique 3) Évaluer et gérer les risques pour la continuité numérique 4) Maintenir la continuité numérique.

NCDCR Digital Preservation Best Practices and Guidelines - Créer des fichiers numériques
http://digitalpreservation.ncdcr.gov/
Lancé pour la première fois en 2010 par le *Digital Information Management Program*, la *State Library of North Carolina* et la *Digital Services Section of the State Archives of North Carolina*, ce site a reçu un prix de l'innovation de la *National Digital Stewardship Alliance* en 2012. L'objectif est de fournir des informations pratiques et introductives sur la préservation numérique et d'orienter les visiteurs vers des ressources accessibles sur la prochaine étape.

Digital Preservation Management Tools and Techniques
http://dpworkshop.org/workshops/management-tools
Les livrables décrivant les processus de travail, issus des ateliers *DPM*) comprennent un diagramme de haut niveau et des diagrammes de plus bas niveau pour la gestion du contenu physique, la transition vers la numérisation et la gestion du contenu numérique natif et numérisé. L'idée est de fournir un processus de travail commun pour tous les contenus dans n'importe quel contexte, puis de développer un certain nombre de cas d'utilisation pour mettre en évidence les exceptions pour des types de contenus spécifiques avec différents types d'exigences.

[37] *Data Protection Act.*
[38] *Freedom of Information Act.*
[39] *Statistics and Registration Services Act.*

Première partie : Pourquoi le nom des fichiers est-il important ?
https://www.youtube.com/watch?v=Hi_A4Ywn4VU
Cette excellente courte vidéo constitue la première des quatre parties d'un tutoriel sur le nommage des fichiers. Elle explique pourquoi il est important de choisir judicieusement les noms de vos fichiers. Conçue pour le grand public, elle fait partie de la série *Inform U* de la Bibliothèque d'État de Caroline du Nord (2012, 3mins 19 sec).

Études de cas

Note de cas de la DPC : le *University of London Computer Center*(ULCC) évalue l'accès à long terme à partir de projets de numérisation à court terme
http://www.dpconline.org/component/docman/doc_download/534-casenoteassessingpreservationindigitization.pdf
Les projets de numérisation sont pour la plupart financés à court terme, alors comment pouvons-nous prendre des mesures pour que les résultats de la numérisation soient solides à long terme ? Cette étude de cas, financée par le *Jisc*, rend compte des travaux entrepris par le *University of London Computer Centre*[40] pour évaluer les plans à long terme de 16 projets de numérisation. Elle fournit un outil d'enquête de base pour aider les financeurs et les responsables de projets à réfléchir à leurs plans de préservation à long terme. Novembre 2010 (4 pages).

Projet *Save our Sounds* de la *British Library*
http://www.bl.uk/projects/save-our-sounds
Lancé en 2015, *Save our Sounds* est le programme de la *British Library* visant à préserver par la numérisation les archives sonores de la nation, une collection de plus de 6,5 millions d'enregistrements de discours, de musique, de la faune et de l'environnement, des années 1880 à nos jours. Le projet vise à la fois à garantir que les archives existantes sont correctement préservées et qu'il existe des systèmes adéquats pour l'acquisition de la future production sonore au Royaume-Uni.

Études de cas du *Digital Curation Center*
http://www.dcc.ac.uk/resources
En 2013, le DCC a commencé une série d'études de cas pour accompagner le nouveau guide du DCC *: "How to Develop Research Data Management Services?".* Celles-ci couvrent des éléments spécifiques d'un service de gestion des données de recherche intéressant les chercheurs et les gestionnaires de données.

[40]Le Centre informatique de l'Université de Londres.

Études de cas sur le campus de la *Society of American Archivists*

http://www2.archivists.org/publications/epubs/Campus-Case-Studies

Les *Campus Case Studies* sont des rapports d'archivistes universitaires américains qui ont créé des solutions de travail. Ils couvrent un large éventail de sujets dont certains sont spécifiquement axés sur la préservation numérique et la création de documents numériques. L'actualité des études de cas varie de 2008 à aujourd'hui.

Why metadata matters

https://cbaileymsls.wordpress.com/2013/09/29/metadata/

Cet article de blog donne de bons exemples pour expliquer pourquoi un nom de fichier et une description des métadonnées médiocres lors de la création d'un fichier peuvent entraver la recherche, la découverte et la réutilisation ultérieures.

Références

Archives nationales du Royaume-Uni, 2015. *Digitisation at The National Archives*. Disponible à l'adresse suivante : http://nationalarchives.gov.uk/documents/information-management/digitisation-at-the-national-archives.pdf

Waters, D and Garrett, J., 1996. *Preserving Digital Information: Report of the Task Force on Archiving of Digital Information commissioned by the Commission on Preservation and Access and the Research Libraries Group*. Washington, DC: *Commission on Preservation and Access*. Disponible à l'adresse suivante : https://www.oclc.org/content/dam/research/activities/digpresstudy/final-report.pdf

Acquisition et évaluation

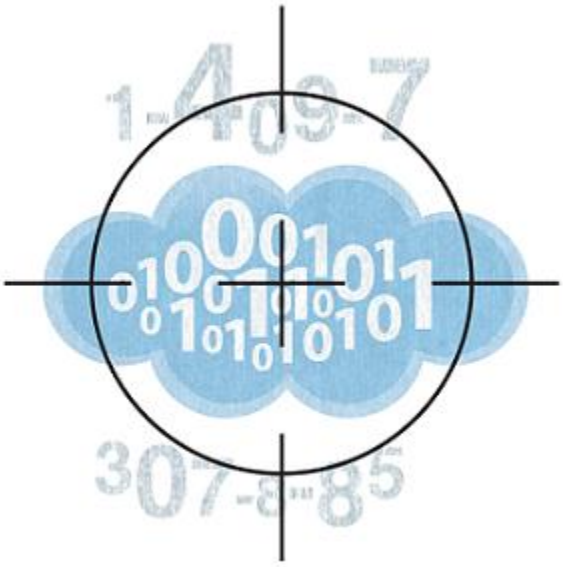


Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Introduction

Dans un environnement numérique, les décisions prises en matière de création et de sélection ont des implications importantes pour la préservation. Le lien entre l'accès et la préservation est beaucoup plus explicite que pour le papier et les autres matériaux traditionnels, car l'accès à une ressource numérique peut être perdu dans un laps de temps relativement court si des mesures actives ne sont pas prises pour la maintenir (c'est-à-dire la préserver) dès le début. Comme l'indique l'arbre de décision interactif, s'il n'est ni possible ni souhaitable de préserver une ressource numérique à travers les divers changements technologiques, il peut être nécessaire de réévaluer son acquisition. Si bon nombre des mêmes principes de l'environnement de préservation traditionnel peuvent être utilement appliqués, les politiques et procédures devront être adaptées à l'environnement numérique.

Dans un environnement imprimé, les dimensions physiques d'une archive signifient que les raisons de la sélection sont relativement bien comprises, tandis que la décision de préserver peut être prise de manière tout à fait distincte et dans un délai qui peut s'étendre sur plusieurs décennies. En revanche, les ressources numériques peuvent s'accumuler et l'évaluation peut être une tâche ardue. En outre, elles peuvent devenir inaccessibles relativement rapidement, de sorte qu'il devienne nécessaire de paralléliser les décisions relatives à la sélection et à la préservation pour les collections numériques.

Si cela induit qu'une plus grande rigueur est nécessaire dans la sélection des ressources numériques que pour les documents imprimés ou autres documents analogiques, cela permettra d'éviter des coûts supplémentaires ultérieurs, car la préservation rétrospective des ressources numériques n'est pas recommandée.

Une documentation précise est également cruciale dans l'environnement numérique. Elle fournira non seulement les détails essentiels pour gérer la ressource dans le temps, mais aussi des informations sur le contexte sans lesquelles il ne sert pas à grand-chose de préserver l'objet numérique lui-même, même si cela est techniquement possible. Dans l'arbre de décision ci-joint, il est suggéré de réévaluer l'acquisition si la documentation est inadéquate.

Dans le cas des ressources numériques en réseau, où l'accès à une ressource ne nécessite pas nécessairement l'introduction physique de la ressource dans une collection, le concept d'acquisition est très différent des collections traditionnelles. Il existe toute une série d'options pour fournir l'accès ou constituer des « collections virtuelles » : par exemple, faire des copies/miroirs pour l'accès, fournir un hyperlien vers une ressource, des catalogues en ligne et des instruments de recherche.
Dans certains cas, une institution peut être réticente à assumer la responsabilité première de la préservation d'un objet si elle estime que l'intérêt pour sa préservation est unanimement reconnu de telle sorte qu'il constituerait une charge injuste pour sa propre institution. Cela souligne la nécessité d'une collaboration entre les institutions et la nécessité d'établir des accords équitables pour des efforts partagés lorsque cela est nécessaire. Un certain nombre de services ont vu le jour ces dernières années, comme le registre des détenteurs de périodiques électroniques – engagés pour préserver des objets qui peuvent être d'intérêt général. L'arbre de décision pour l'évaluation et la sélection qui l'accompagne est basé sur l'hypothèse que la ressource n'a pas encore été acquise et indique un certain nombre de points pour lesquels il faudra tenir compte des implications financières avant de prendre la décision de procéder à l'acquisition. Il suggère qu'à ces points, des choix difficiles devront peut-être être faits pour déterminer si la ressource justifie les coûts ou s'il est préférable de ne pas procéder à l'acquisition.

Développement de la politique et de la procédure

Avant de se lancer dans l'acquisition et la prise en charge de collections numériques, il peut être nécessaire de déterminer si les politiques (par exemple, le développement des collections) et les procédures actuelles sont toujours adaptées à l'ère numérique (voir Politiques et stratégies institutionnelles). Selon la structure et la formulation des documents existants, cet examen peut aboutir à des changements mineurs qui élargissent le champ d'application pour inclure les objets numériques, ou à la création de nouveaux documents de politique générale couvrant spécifiquement les collections numériques. Les ajouts ou modifications à la politique peuvent inclure des descriptions des types d'objets qui seront acquis, en relation avec le format et/ou le contenu, ainsi que l'examen d'autres questions telles que les droits de propriété intellectuelle, la sensibilité et les considérations d'accès. Il est essentiel que toute modification soit ratifiée par les comités de gestion compétents au sein de votre organisation afin de garantir le soutien et l'adhésion.

Politique d'évaluation/rétention

L'arbre de décision qui accompagne cette section peut être utilisé comme un outil pour construire ou tester la politique de sélection, d'évaluation ou de rétention de votre organisation.
L'évaluation des objets nativement numériques doit inclure une évaluation mesurée de leur valeur pour l'organisation mère, en fonction des défis que représentent la préservation à long terme et la fourniture d'un accès. Ces défis peuvent comprendre la capacité d'une organisation à lire ou à ouvrir une version du fichier maître, la capacité à obtenir des droits suffisants pour gérer et fournir l'accès aux versions actuelles et futures du fichier, ou simplement les ressources en personnel et en financement. Les organisations doivent donc se concentrer dans un premier temps sur un équilibre entre l'acquisition d'objets numériques de grande valeur et ces obligations de préservation à long terme. Il convient de rappeler que les organisations peuvent donner accès aux ressources auxquelles elles ont accès sans les placer dans des processus de travail spécifiques de préservation ou de préservation. Un document politique détaillé qui identifie clairement les ressources numériques les plus importantes (du point de vue du format ou du contenu) peut donner des indications sur l'évaluation des objets nativement numériques destinés à ces filières. Pour les acquisitions numériques de moindre valeur, qui sont souvent associées à des acquisitions

de plus grande valeur, il peut suffire de décrire le niveau d'accès et de préservation qu'une organisation leur fournira. Cette description doit inclure l'indication d'un calendrier de préservation adapté à ce type de contenu. Cela peut signifier également l'inclusion d'un calendrier de mise au rebut ou d'une politique de désaffectation, le cas échéant (voir Préservation et (ré)évaluation).

Accords et orientations pour les déposants – formats de fichiers, documentation requise

Une fois la politique établie, un certain nombre de pièces justificatives supplémentaires seront nécessaires pour faciliter le processus d'acquisition et d'évaluation. Outre les documents de procédure standard, une organisation peut souhaiter créer une série d'accords et de licences standard pour les déposants afin de faciliter le processus de négociation. Ces documents seront particulièrement utiles pour garantir l'octroi des autorisations et des droits de propriété intellectuelle minimums requis pour la préservation. Sans accords de licence suffisants, une organisation peut se retrouver en possession de collections numériques qu'elle ne détient pas les droits de préserver activement ou auxquelles elle ne donne pas accès (voir Conformité juridique). Ces dispositions peuvent également être complétées par des notes d'orientation à l'intention des déposants qui définissent les exigences relatives aux objets à transférer et aux documents d'accompagnement.

Normes d'acquisition et de transfert

L'expérience montre que le transfert d'un producteur à une Archive peut être tortueux et que, par conséquent, tout outil susceptible de rationaliser le processus est susceptible de profiter aux deux parties. Deux initiatives ont tenté de normaliser l'interface entre les producteurs et les archives en un processus cohérent et clair, en cultivant une compréhension mutuelle entre les producteurs et les archives en ce qui concerne leurs rôles respectifs : Interface producteur-Archives – Norme de résumé de méthodologie (PAIMAS, ISO 20652:2006) ; et la norme d'interface producteur-Archives (PAIS, ISO 20104:2015).

PAIMAS fournit une description normalisée des interactions entre les producteurs et une Archive. Il segmente le processus de transfert en un certain nombre de phases, en fournissant une description détaillée du résultat attendu de chaque phase et des actions requises pour obtenir ce résultat. Les quatre phases principales – phase préliminaire, phase de définition formelle, phase de transfert, phase de validation – servent de base pour identifier les domaines de l'interface producteur-Archive qui bénéficieraient de normes, de recommandations et de meilleures pratiques plus ciblées, et fournissent également une base pour le développement de processus automatisés et d'outils logiciels pour soutenir le processus de transfert d'informations. PAIMAS étend implicitement les exigences détaillées pour l'Entrée et l'Administration dans le modèle de référence OAIS.
PAIS fournit une méthode normalisée pour définir formellement les objets d'information numérique à transférer par un producteur d'information vers un entrepôt numérique et pour empaqueter efficacement ces objets sous la forme de paquets d'informations à verser. Il vise à fournir des définitions plus précises des objets numériques, en aidant l'Archive à traiter et à valider les objets reçus lors de la soumission.

Processus d'acquisition

Négociation

La négociation des conditions du dépôt doit avoir lieu avant le transfert des documents. De nombreux aspects de l'accord de dépôt peuvent être couverts par la politique d'acquisition de

l'organisation, mais des détails sur chaque dépôt, en particulier pour les archives locales et spécialisées, peuvent être requis au niveau de la collection. Le déposant doit indiquer s'il existe des limitations quant à la nature et au moment de la publication des documents. Par exemple, certains documents peuvent-ils être ouverts immédiatement, tandis que d'autres ne peuvent être ouverts qu'au décès du déposant ou après une période déterminée ?
Une considération essentielle est le droit de l'organisation de modifier le document à des fins de préservation, par exemple la migration vers un format qui peut être conservé à long terme ou accessible.
Si le transfert comprend un contenu essentiel à la compréhension des documents mais ne constitue pas un document en soi, il doit y avoir un accord selon lequel l'organisme peut supprimer ces fichiers lorsque leur contenu a été saisi pour être utilisé ailleurs (par exemple comme métadonnées pour les documents).

Transfert

La plupart des institutions devront élaborer des procédures et des documents pour favoriser le transfert sans heurts des ressources numériques des fournisseurs vers leurs collections.
Lors du transfert d'objets numériques nés dans l'environnement informatique d'une organisation, il convient de réfléchir à la manière dont cela doit se faire afin de garantir la sécurité et l'exhaustivité du transfert. Pour les petites organisations, il peut être suffisant de faire livrer les fichiers numériques pertinents sur un lecteur ou un matériel similaire et de vérifier leur contenu par rapport au manifeste descriptif du fichier. Les organisations peuvent également souhaiter transférer les fichiers numériques en utilisant un FTP interne, ou une solution payante pour un tiers (comme un service de partage de fichiers dans le *cloud*), afin de garantir la chaîne de préservation.
Le tableau ci-dessous présente les options de transfert et de processus d'entrée des objets numériques. Les décisions relatives aux formats de fichiers et, le cas échéant, aux supports de stockage (voir Stockage, Anciens supports et Formats de fichiers et normes) prendront en charge ce processus et en seront interdépendantes.

Options pour le transfert et l'entrée des formats de fichiers et des supports de stockage		
Options	**Points d'attention**	**Exigences**
Toutes les options		• Politique sur les formats de fichiers et les normes. • Planification de la préservation et veille technologique sur les développements en matière de stockage et de formats de supports hérités.
Limiter l'éventail des formats de fichiers reçus **Limiter la gamme de médias reçus (option la plus rentable à long terme)**	• Simplifie la gestion et réduit les coûts globaux. • Le déposant peut manquer de ressources ou d'expertise pour se conformer. • Grande variété de formats de fichiers utilisés et d'extensions propriétaires aux normes ouvertes.	• Directives sur les formats de fichiers préférés. • Degré d'influence sur le dépôt. • Stratégies de sensibilisation et de collaboration pour atteindre les résultats souhaités. • Lignes directrices sur les

	• Les supports de stockage physique utilisés pour le transfert peuvent n'être que des supports temporaires et le contenu sera transféré vers le stockage à long terme utilisé.	supports de transfert préférés et les procédures de transfert.
Accepter les formats de fichiers tels qu'ils sont reçus mais les convertir en format de fichier standard Accepter les supports de stockage tels qu'ils sont reçus mais transférer les contenus sur le stockage standard utilisé	• Simplifie la gestion et réduit les coûts à long terme. • Peut ne pas être techniquement réalisable pour la conversion au format de fichier standard. • Il sera nécessaire de vérifier qu'il n'y a pas eu de perte accidentelle de données.	• Respect de la législation, autorisations de droits d'auteur ou droits de préservation statutaires. • Ressources et expertise technique de l'établissement d'accueil. • Choix des formats préférés. • Documentation des formats natifs pour permettre la conversion. • Contrôles d'intégrité pour le processus de conversion.
Accepter et conserver tel quel (option la moins rentable à long terme, malgré des coûts initiaux plus faibles)	• Complique la gestion et augmente les coûts de gestion des ressources au fil du temps. • Option à haut risque, en particulier si un grand nombre de documents numériques sont collectés. • Un choix de formats de fichiers peut être disponible. Celui qui est déposé peut ne pas être le plus adapté à la préservation. • Les supports de stockage peuvent être de qualité inconnue et ne pas convenir à la préservation à long terme. • Les formats peuvent être obsolètes ou ne pas être pris en charge par l'institution.	• Des priorités clairement définies pour la préservation à court et à long terme. • Capacité à traiter des questions telles que le cryptage, les logiciels propriétaires, etc. dans les articles reçus. • Capacité à garantir l'accès futur aux informations contenues dans l'article.

Validation

Une fois le transfert effectué, les fichiers doivent être placés dans un environnement sécurisé, en quarantaine et sauvegardé, et un processus d'archivage doit être rapidement mis en place. Après avoir transféré et hébergé une copie d'une collection nativement numérique, une organisation est désormais chargée de certaines responsabilités légales telles que les demandes d'accès à l'information si elle fait partie du secteur public. Par la suite, un accusé de réception doit être envoyé au donateur. Il est important à ce stade précoce qu'aucune instruction de destruction des fichiers originaux ne soit donnée.

Les dossiers doivent faire l'objet d'un contrôle anti-virus dès que possible afin de s'assurer que l'objet n'a pas été infecté par un logiciel malveillant ou un virus. Si certains sont détectés, l'organisme déposant doit être alerté et les supports doivent lui être retournés (s'il n'en a pas de copie) ou être formatés et retournés ou détruits selon la préférence du déposant. Une fois qu'il est confirmé que les documents sont exempts de virus, il convient de vérifier que tous les documents sont présents et intacts. La méthode la plus fiable pour cela est de vérifier les fichiers par rapport au manifeste. Créez des empreintes pour les fichiers et comparez-les à celles qui figurent sur le manifeste avant le transfert. Si les empreintes correspondent, vous pouvez être sûr que les dossiers n'ont pas été corrompus ou altérés accidentellement entre les points de transfert du déposant et l'arrivée à l'organisation. Si aucun manifeste vérifiable n'a été fourni avec le dépôt, il peut être impossible de vérifier de manière exhaustive l'intégrité des dossiers et la consultation manuelle d'un échantillon de dossiers peut être nécessaire pour donner une indication de l'exhaustivité et de la qualité. Dans ce cas, un manifeste vérifiable doit être généré pour permettre une vérification ultérieure de la fixité.

À ce stade, on peut espérer avoir la certitude que les fichiers sont complets (selon le manifeste), qu'ils conservent leur intégrité (exactement ce que le déposant a fourni) et qu'ils sont exempts de virus et de logiciels malveillants. Ils peuvent maintenant être pris en charge dans le système de préservation numérique.

Les métadonnées décrivant l'objet déposé aideront à garantir la fixité (voir Intégrité et empreintes) de l'objet pendant le processus de transfert ainsi qu'à soutenir la préservation et l'accès ultérieurs. Il peut s'agir notamment des éléments suivants :

- Un manifeste vérifiable comprenant une liste des noms de fichiers et de dossiers et empreintes/valeurs de fixité pour chaque fichier
- La taille des dossiers (avec un volume total)
- Une liste des formats de fichiers
- Une déclaration détaillant tout DPI associé aux enregistrements

Dans la mesure du possible, c'est au déposant qu'il incombe de fournir des informations sur les droits de propriété intellectuelle des documents.

Processus de prise en charge

La période entre le transfert à l'organisation et l'intégration dans le dépôt de l'organisation ou l'environnement de préservation numérique peut être importante. Cette phase d'acquisition peut être particulièrement longue pour les grandes collections numériques naissantes, s'étendant parfois sur plusieurs années, mais c'est au cours de cette phase qu'une évaluation qualitative des objets peut être effectuée. Les objets sont examinés, leurs métadonnées techniques sont recueillies, leurs métadonnées descriptives sont améliorées et les processus généraux d'entrée de l'organisation applicables à tout objet prennent le relais.

C'est au cours de cette période d'évaluation parfois prolongée que la prise en charge des objets peut être réévaluée, ou rejetée si, à l'examen, on estime que les objets ne correspondent pas au profil d'acquisition ou de collection de l'organisation, à la spécification du format de fichier établie dans les documents d'orientation, ou pour toute autre raison. Un moratoire peut être imposé sur des éléments particulièrement sensibles tels que les informations personnelles, les informations commerciales sensibles ou les éléments qui enfreignent les lois sur la diffamation, par exemple. Dans de tels cas, il est important de préciser clairement la période de fermeture du dossier.

Procédures de préparation des données et de la documentation en vue de leur stockage et de leur préservation
Numérotation unique Chaque ressource numérique acquise par une institution doit se voir attribuer un identifiant unique. Ce numéro identifiera la ressource dans le catalogue de l'institution et sera utilisé pour localiser ou identifier les supports physiques et la documentation. En cas de désaffectation d'une ressource pour une raison quelconque, ce numéro unique ne doit pas être réattribué. Consultez la rubrique Identifiants pérennes pour obtenir des conseils si vous utilisez un schéma d'identifiants persistants.
Lignes directrices pour les transferts de manutention et de stockage Il convient d'élaborer des directives de manipulation et de transfert à l'intention du personnel en place, en tenant compte des conseils des personnels des services informatiques et de préservation sur les meilleures pratiques pour les différents supports de stockage et le transfert de fichiers vers des systèmes de stockage à long terme (voir Anciens supports, Criminalistique numérique et Stockage).
Reformatage des formats de fichiers Lorsque les formats de fichier utilisés pour transférer la ressource ne conviennent pas à la préservation à long terme, l'institution peut reformater la ressource dans ses formats de fichier préférés. En plus des formats d'archives, des versions dans d'autres formats adaptés à la délivrance aux utilisateurs peuvent également être produites à partir de l'original (voir Formats de fichiers et normes, et Stockage).
Copie Plusieurs copies de sauvegarde d'un article peuvent être générées lors du processus d'entrée dans le cadre de la politique de stockage et de préservation des institutions et pour permettre des procédures de reprise après sinistre (voir Stockage)
Sécurité Des politiques et des procédures de sécurité physique et du système doivent être mises en place pour garantir le soin et l'intégrité des articles pendant le processus d'entrée. Ces politiques et procédures doivent être élaborées à partir des politiques et procédures institutionnelles en matière de sécurité et en tenir compte (voir Sécurité de l'information).
Édition et contrôle des versions Procédures de mise à jour et de contrôle de l'édition de tout objet numérique dynamique auquel on a accédé (par exemple, des instantanés annuels des bases de données qui sont régulièrement mises à jour) ou de contrôle de la version des éléments auxquels on a accédé, le cas échéant (par exemple, des éléments auxquels on a accédé dans différents formats ou pour lesquels différents formats de préservation et d'accès ont été générés).
Normes de catalogage et de documentation Les métadonnées et la documentation reçues ou créées pendant le transfert, la validation et la prise en charge sont essentielles pour échanger efficacement des informations et des documents entre les plateformes et les personnes. Au minimum, elles doivent fournir des informations sur la provenance et l'historique administratif d'un objet (y compris tout traitement

de données effectué depuis sa création), son contenu, sa structure, ainsi que sur les modalités et conditions liées à sa gestion et à son utilisation ultérieures, y compris les droits de propriété intellectuelle et la période à laquelle ils s'appliquent (voir Métadonnées et documentation). Elles doivent être suffisamment détaillées pour décrire : • La découverte de la ressource (par exemple, l'emplacement d'une ressource qui est au moins brièvement décrite avec de nombreuses autres ressources) • L'évaluation de la ressource (par exemple, le processus par lequel un utilisateur détermine s'il a besoin d'accéder à cette ressource). • La commande de ressources (par exemple, les informations qui renseignent l'utilisateur sur les conditions d'utilisation d'une ressource et les processus ou autres moyens d'accès à cette ressource). • L'utilisation de la ressource (par exemple, les informations dont un utilisateur peut avoir besoin pour accéder au contenu informatif de la ressource). • La gestion des ressources (par exemple, les informations administratives essentielles à la gestion et à la préservation d'une ressource dans le cadre d'une collection plus large et comprenant des informations sur l'emplacement, le contrôle des versions, etc.)
Temps de traitement Idéalement, des objectifs devraient être fixés et contrôlés pendant la période maximale entre l'acquisition et le catalogage afin d'éviter l'accumulation d'objets non traités et potentiellement à risque au cours du processus d'entrée.

Compétences, ressources et capacités

Les organisations doivent examiner si elles disposent de ressources techniques et humaines suffisantes pour acquérir des collections numériques. Cette information peut cependant ne pas être visible au début d'une acquisition, car les divers défis de la préservation de collections numériques spécifiques peuvent ne se révéler qu'avec le temps. Les organisations doivent donc prévoir les lacunes en matière de connaissances, de compétences et de personnel et, si possible, y remédier par la formation, le recrutement ou l'engagement de services professionnels spécialisés dans la préservation numérique. Lorsque les ressources financières ne permettent pas de les combler, une campagne interne de renforcement des connaissances peut souvent suffire pour l'intérim. (voir Formation et développement des compétences du personnel, et Achats et recours à de tiers).

Coûts d'acquisition et de prise en charge

Essayer d'établir des coûts indicatifs pour les activités de préservation numérique est toujours problématique. Ceux-ci ne devraient pas seulement inclure le stockage (le plus évident) mais aussi le coût des ressources humaines nécessaires pour gérer l'acquisition et la prise en charge de chaque objet nativement numérique, un processus qui peut faire écho aux modalités d'acquisition des objets physiques. D'autres coûts anticipés pourraient inclure des processus de préservation tels que la normalisation, l'analyse, l'enrichissement des métadonnées, la robustesse accrue du stockage, la récupération après sinistre, etc.

Bien qu'une organisation doive trouver les solutions les plus rentables pour ces coûts de cycle de vie, il faut reconnaître que l'investissement nécessaire pour fournir une voie de préservation robuste qui puisse sauvegarder notre patrimoine numérique peut être important, et que certains processus doivent être respectés quelle que soit la nature de l'objet acquis. Les organisations

doivent donc garder cela à l'esprit lorsqu'elles acquièrent des collections numériques naissantes (voir Analyses de rentabilité, avantages, coûts et impact).

Résumé des recommandations

Acquisition et évaluation – liste de contrôle des recommandations

Accords et orientations pour les déposants

- ☐ Créer une série d'accords et de licences types pour les déposants
- ☐ Élaborer des conseils appropriés pour les déposants

Procédures de transfert

- ☐ Fournir de la documentation pour guider et soutenir le transfert d'objets numériques depuis les fournisseurs
- ☐ Décider comment vos procédures de transfert peuvent être développées au mieux pour soutenir vos politiques de stockage et de préservation

Procédures de validation

- ☐ Vérifier les médias, le contenu et la structure

Procédures de préparation des données et de la documentation en vue de leur stockage et de leur préservation

- ☐ Numérotation unique de chaque article déposé
- ☐ Directives de manipulation et de transfert de stockage pour les différents supports
- ☐ Reformatage des formats de fichiers si nécessaire selon les lignes directrices convenues
- ☐ Générer des copies multiples d'un article dans le cadre de la politique de stockage et de préservation d'une institution
- ☐ Politique et procédures de sécurité des systèmes et des biens physiques pendant le processus d'entrée

Procédures de catalogage et de documentation

- ☐ Une norme minimale d'information requise pour le catalogage, y compris les informations sur les DPI
- ☐ Directives pour la documentation rétrospective ou l'enrichissement du catalogue.
- ☐ Procédures de mise à jour et de gestion des versions ou éditions d'un article.
- ☐ Procédures de mise à jour des bases de données de gestion des collections
- ☐ Sélection des normes de catalogage et de documentation
- ☐ Objectifs des tâches liées au processus d'entrée et calendrier de leur réalisation

Révision des procédures

- ☐ Les lignes directrices et les calendriers devraient idéalement être révisés chaque année, ou aussi souvent que possible pour suivre l'évolution des besoins et des politiques de développement des collections d'une organisation

Formation des équipes

- ☐ Prévoir les lacunes en matière de connaissances, de compétences et de personnel et, si possible, y remédier par la formation, le recrutement ou le recours à des services spécialisés de tiers

Coûts

☐ Évaluer et planifier les coûts du cycle de vie des acquisitions

Ressources

ISO 20104:2015 - Systèmes de transfert de données et d'informations spatiales – Spécification d'interface producteur-Archive (PAIS)

CCSDS 651.1-B-1, Spécification d'interface producteur-Archive (PAIS,) Norme recommandée, CCSDS 651.1-B-1 livre bleu, Février 2014
https://public.ccsds.org/Pubs/651x1b1.pdf
Le Livre bleu est une pré-impression gratuite de la norme ISO 20104:2015. La norme PAIS vise à fournir une méthode normalisée pour définir formellement les objets d'information numérique à transférer par un producteur d'information vers une Archive et pour empaqueter efficacement ces objets sous la forme de paquets d'information à verser (SIP). Cela permet un transfert et une validation efficaces des données SIP (104 pages).

Version française de la norme ISO 20104:2015 disponible à l'adresse suivante : https://www.iso.org/fr/standard/67056.html

What is appraisal?
http://www.nationalarchives.gov.uk/documents/information-management/what-is-appraisal.pdf
Ces directives des Archives nationales du Royaume-Uni s'appliquent aux documents publics britanniques, quel que soit leur format, y compris le papier, le numérique, l'audio, le film ou le modèle, tels que définis par la loi de 1958 sur les documents publics, et à tous les organismes responsables de ces documents (2013, 7 pages).

Preserving eBooks, DPC Technology Watch Report, 14-01 July 2014
http://dx.doi.org/10.7207/twr14-01
Ce rapport examine les développements actuels et les questions avec lesquelles les bibliothèques publiques, nationales et d'enseignement supérieur, les éditeurs, les agrégateurs et les institutions de préservation doivent composer pour assurer un accès à long terme au contenu des livres électroniques et qui affectent l'acquisition aussi bien que la préservation (31 pages).

Preservation, Trust and Continuing Access for e-Journals, DPC Technology Watch Report 13-04 September 2013
http://dx.doi.org/10.7207/twr13-04
Ce rapport examine les développements et les problèmes actuels auxquels sont confrontés les bibliothèques, les éditeurs, les intermédiaires et les fournisseurs de services dans le domaine de la préservation numérique, de la confiance et de l'accès continu aux périodiques électroniques. Il n'est pas uniquement axé sur la technologie et couvre les questions juridiques, économiques et de service pertinentes dans l'acquisition de l'accès aux ressources numériques en réseau et les défis uniques de préservation que cela représente (43 pages).

The UNESCO/PERSIST Guidelines for the selection of digital heritage for long-term preservation
https://unesdoc.unesco.org/ark:/48223/pf0000244280
Le projet UNESCO/PERSIST (Plate-forme pour le renforcement de la durabilité de la société de l'information à l'échelle mondiale) a publié ces lignes directrices sur la sélection du patrimoine

numérique pour la préservation à long terme en mars 2016. L'objectif de ces lignes directrices est de fournir un point de départ général aux bibliothèques, archives, musées et autres institutions patrimoniales pour l'élaboration de leurs propres politiques de sélection du patrimoine numérique en vue d'une préservation numérique durable à long terme (19 pages).

Community Owned digital Preservation Tool Registry COPTR
http://coptr.digipres.org/Main_Page
Le registre COPTR décrit des outils utiles pour la préservation numérique à long terme et agit principalement comme un outil de recherche et d'évaluation pour aider les professionnels à trouver les outils dont ils ont besoin pour préserver les données numériques. Le registre COPTR vise à rassembler en un seul endroit les connaissances de la communauté de la préservation numérique sur les outils de préservation. Il a été initialement alimenté par des données provenant de registres gérés par les organisations partenaires du registre COPTR, notamment ceux tenus par le *Digital Curation Centre*, le *Digital Curation Exchange, la National Digital Stewardship Alliance*, l'*Open Preservation Foundation* et le *Preserving digital Objects With Restricted Resources project* (POWRR). Le registre COPTR permet de saisir des détails factuels de base sur un outil, ses fonctions, la manière de trouver plus d'informations (URL pertinentes) et des références aux expériences des utilisateurs avec l'outil. Le champ d'application est une interprétation large du terme « préservation numérique ». En d'autres termes, si un outil est utile pour effectuer une fonction de préservation numérique telle que celles décrites dans le modèle OAIS ou le modèle de cycle de vie DCC, alors il entre dans le champ d'application de ce registre.

Keepers Registry
https://keepers.issn.org/
Le registre des détenteurs fait office de contrôleur mondial des modalités d'archivage des périodiques électroniques. Il a trois objectifs principaux : permettre aux bibliothécaires et aux décideurs politiques de savoir qui s'occupe de quel périodique électronique, comment et avec quelles conditions d'accès ; mettre en évidence les périodiques électroniques qui sont encore « en danger de perte » ; et présenter les organisations (les détenteurs) qui font office de rayons numériques pour l'accès à long terme. Il comporte une fonction de comparaison des listes de titres pour vous aider à découvrir le statut d'archivage d'une liste de titres de publications en série importants pour vous : en signalant ceux qui sont archivés et ceux qui sont « à risque ».

MediaRIVERS (*Media Research and Instructional Value Evaluation and Ranking System*)
https://github.com/IUMDPI/MediaSCORE
Un logiciel, créé par l'Université de l'Indiana en collaboration avec *AVPreserve,* qui guide une évaluation structurée de la recherche et de la valeur pédagogique des médias. La version gratuite et open source nécessite une installation et une configuration sur un serveur, et une application hébergée est disponible sur la base d'un abonnement mensuel.

Practical E-Records: software and tools for archivists
http://e-records.chrisprom.com/
Les pages créées par Chris Prom pour les directives de transfert, la politique de dépôt des documents électroniques et le formulaire d'accord de dépôt fournissent des modèles que vous pouvez modifier et/ou fournir aux producteurs de documents dont votre dépôt souhaite l'adhésion.

L'autorisation de modifier et de republier ces directives de transfert est fournie dans le cadre d'une licence Creative Commons Attribution 3.0 United States License.

Archaeology Data Service Guidelines for Depositors
http://archaeologydataservice.ac.uk/advice/guidelinesForDepositors
Les lignes directrices du *Archaology Data Service* (ADS) à l'intention des déposants fournissent des conseils sur la manière de préparer correctement les données et de compiler les métadonnées en vue de leur dépôt auprès du SDA et décrivent les modalités de dépôt des données. Il existe également une série de fiches de travail et de listes de contrôle succinctes couvrant : la gestion des données ; la sélection et la préservation ; les formats de fichier et les métadonnées préférés. D'autres ressources à l'usage des déposants potentiels comprennent une série de guides de bonnes pratiques, qui complètent les lignes directrices du SDA et fournissent des informations plus détaillées sur des types de données spécifiques.

Sélection et transfert des dossiers
http://www.nationalarchives.gov.uk/information-management/manage-information/selection-and-transfer/
Ces pages fournissent des conseils sur la sélection et le transfert des documents. Les organismes britanniques qui transfèrent des documents aux Archives nationales du Royaume-Uni ou à des lieux de dépôt en vertu de la loi de 1958 sur les archives publiques doivent suivre ce processus pour les documents de tous formats et supports, y compris les documents papier et numériques. Il s'agit d'un guide en six étapes :
Étape 1 : Évaluation de vos documents
Étape 2 : Sélection de vos dossiers
Étape 3 : Examen de la sensibilité des documents sélectionnés
Étape 4 : Catalogage et préparation des documents
Étape 5 : Planification et organisation de la livraison des documents
Étape 6 : Accès à vos dossiers

The Work of Appraisal in the Age of Digital Reproduction
http://archival-integration.blogspot.co.uk/2015/06/the-work-of-appraisal-in-age-of-digital.html#pii
La discussion sur le projet d'intégration du processus de travail *ArchivesSpace-Archivematica-DSpace* de la *Bentley Historical Library* met en lumière les techniques actuelles d'évaluation des archives numériques employées par la Bentley, dont beaucoup espèrent les intégrer dans *Archivematica* (juin 2015).

Acquisition et gestion des collections numériques à la Bibliothèque du Congrès
http://www.slideshare.net/NASIG/acquisition-management-of-digital-collections-at-the-library-of-congress-34244613
La Bibliothèque du Congrès, en tant que bibliothèque nationale et siège de l'*US Copyright Office*, est très impliquée dans l'acquisition et la gestion des données numériques. Ce *powerpoint* concis et informatif de Ted Westervelt partage les expériences que la Bibliothèque du Congrès a vécues et

les leçons qu'elle en a tirées (2014, 30 diapositives).«« *Trust Me, I'm an Archivist* » : Expériences avec les donateurs numériques
http://www.ariadne.ac.uk/issue65/hilton-et-al
Cet article publié en 2010 par le personnel de la *Wellcome Trust Library* examine quatre scénarios communs qui semblent constituer de nouveaux obstacles au transfert d'objets numériques : manque de planification à long terme ; TI vs gestion des documents ; duplication et abondance ; et la peur du numérique. Elle conclut que nous devons changer la façon dont nous présentons l'information, dont nous travaillons avec l'objet numérique et dont nous pouvons soutenir et aider nos donateurs. Le degré d'engagement qui est la pratique courante avec les documents papier ne suffira pas pour l'objet numérique natif : notre interaction avec les déposants sera idéalement encore plus étroite et encore plus fréquente, car nous les aidons à faire face non seulement aux nouveaux défis techniques, mais aussi à la pléthore de questions de compétences non techniques, d'idées préconçues et de pièces jointes qui les entourent.

Arbre de décision

Des politiques de sélection clairement définies permettront de réaliser des économies en termes de temps nécessaire pour établir s'il faut ou non sélectionner mais aussi de réduire les coûts potentiels en cas de réévaluation des ressources numériques qui risquent de devenir ou ne sont plus accessibles.

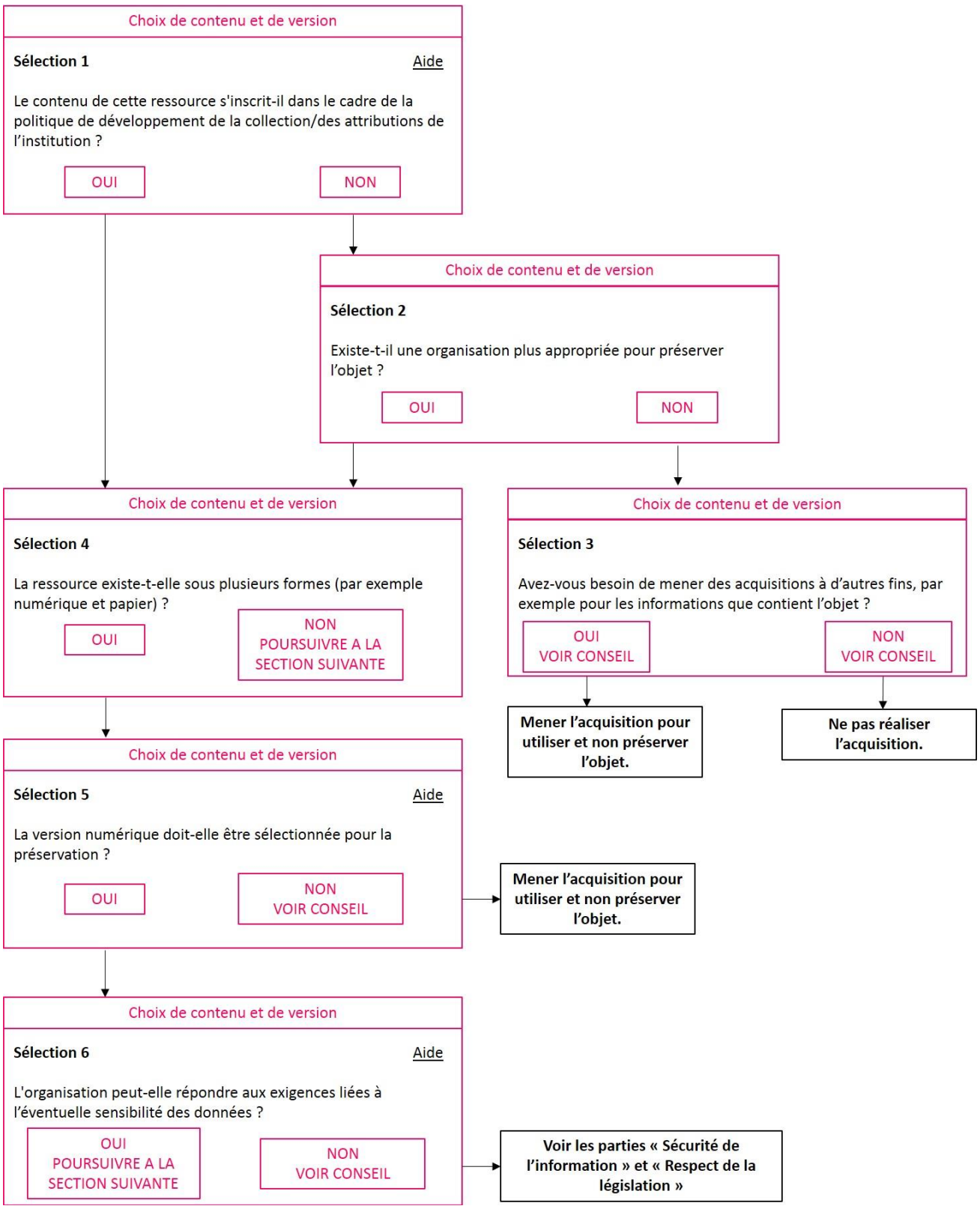
Cet arbre de décision peut être utilisé comme un outil pour construire ou tester une politique de sélection pour votre organisation. Le processus de décision représenté dans l'arbre doit être pris en compte par votre politique de sélection des ressources numériques à long terme.

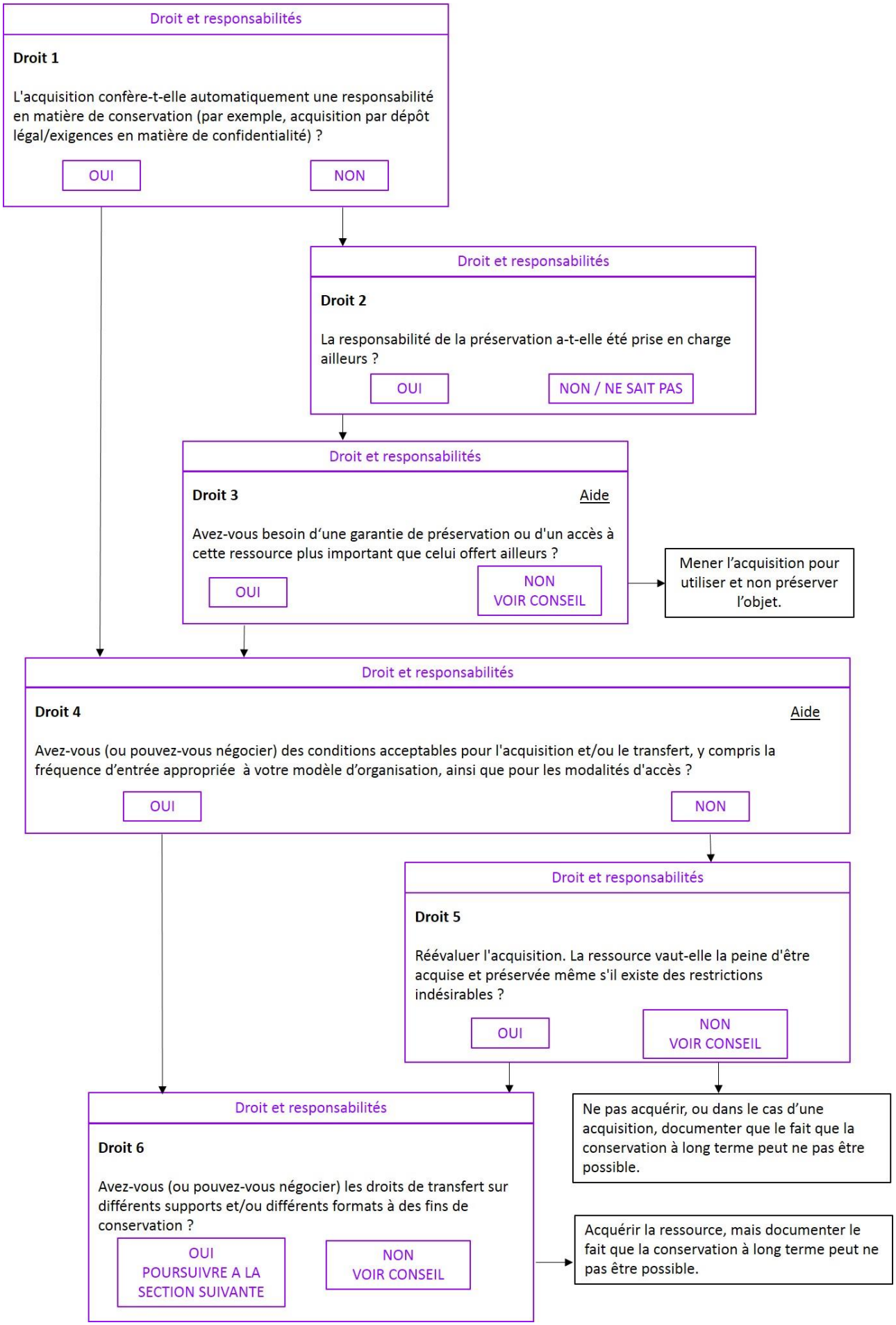
Dans l'hypothèse où une ressource numérique est envisagée pour la sélection, les questions et les choix reflétés ici guideront la décision finale d'accepter ou de rejeter la responsabilité de la préservation à long terme. Le flux des questions représente un ordre logique d'évaluation. Si la réponse aux premières questions n'est pas favorable, il ne sert pas à grand-chose d'accepter la responsabilité de la préservation de la ressource ou de poursuivre son évaluation. Par exemple si le contenu ne répond pas à votre politique de collecte, alors la réponse aux questions sur le format technique ne sera pas pertinente. La structure de l'arbre vise à refléter ce processus.
Rappelez-vous...

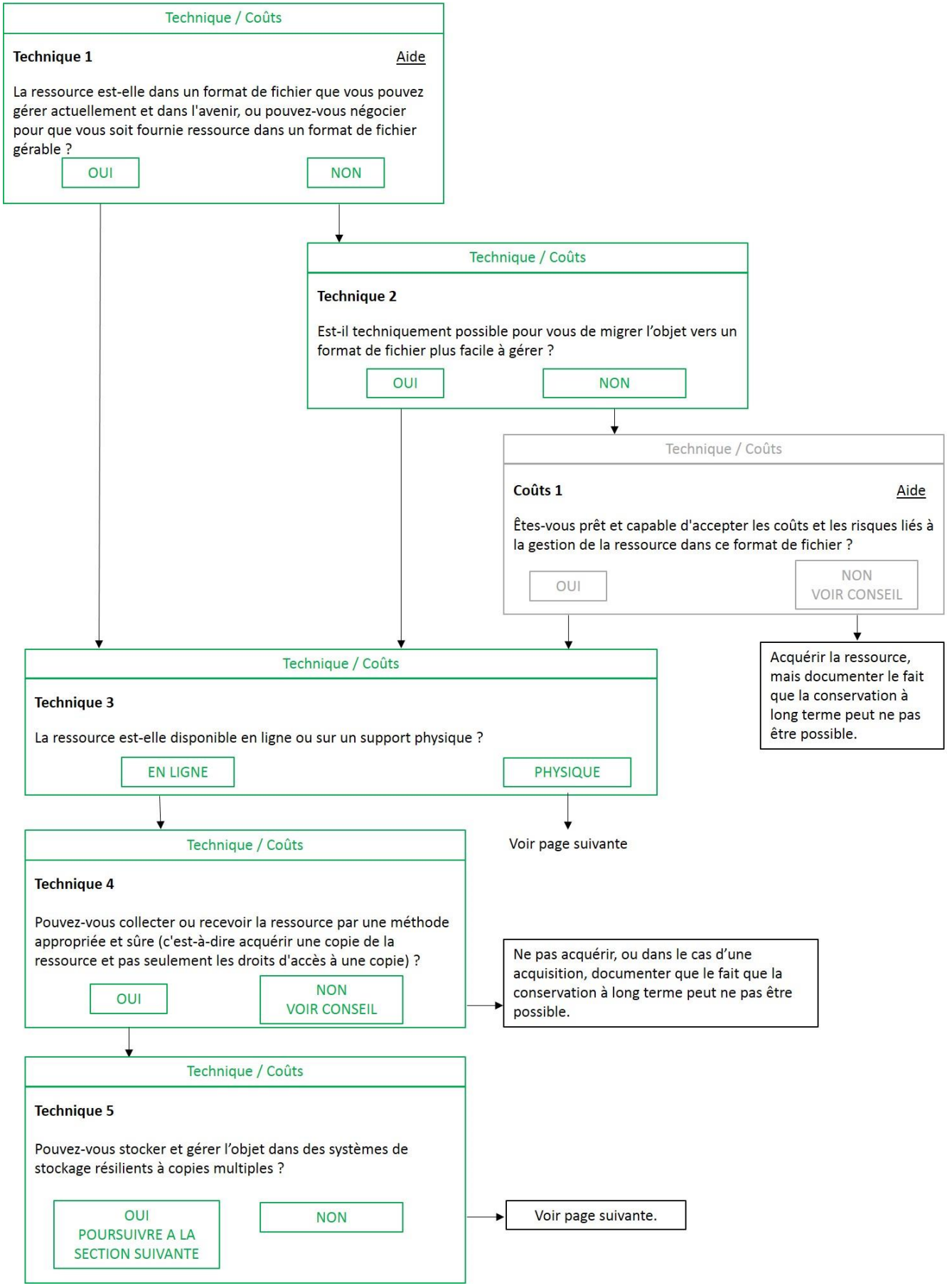
Lorsqu'une politique est en place, pour être efficace, elle doit être aussi :

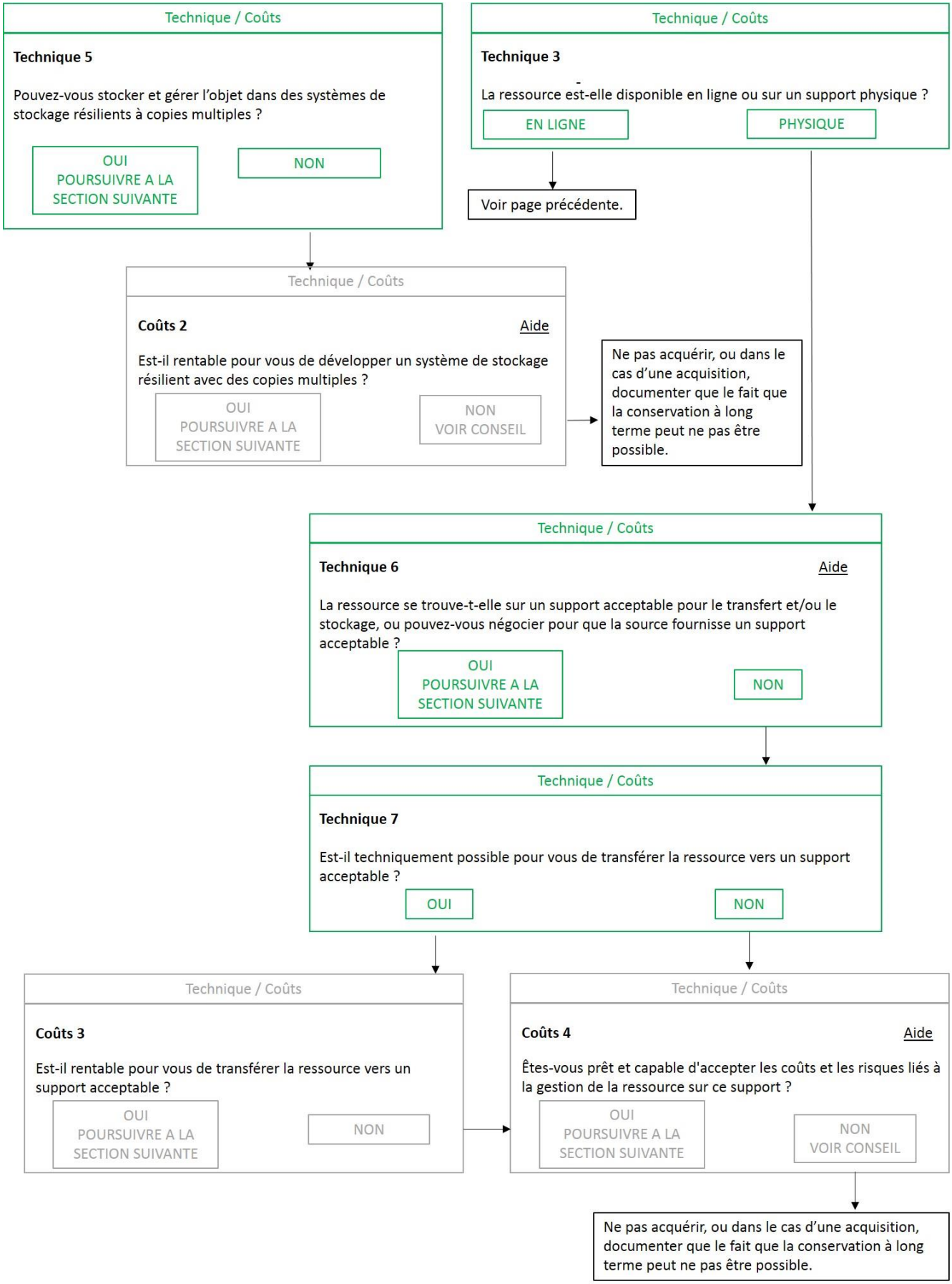
- Approuvée par la direction générale
- Promulguée activement dans toute l'organisation
- Révisée à intervalles réguliers
- Pourvue de ressources appropriées

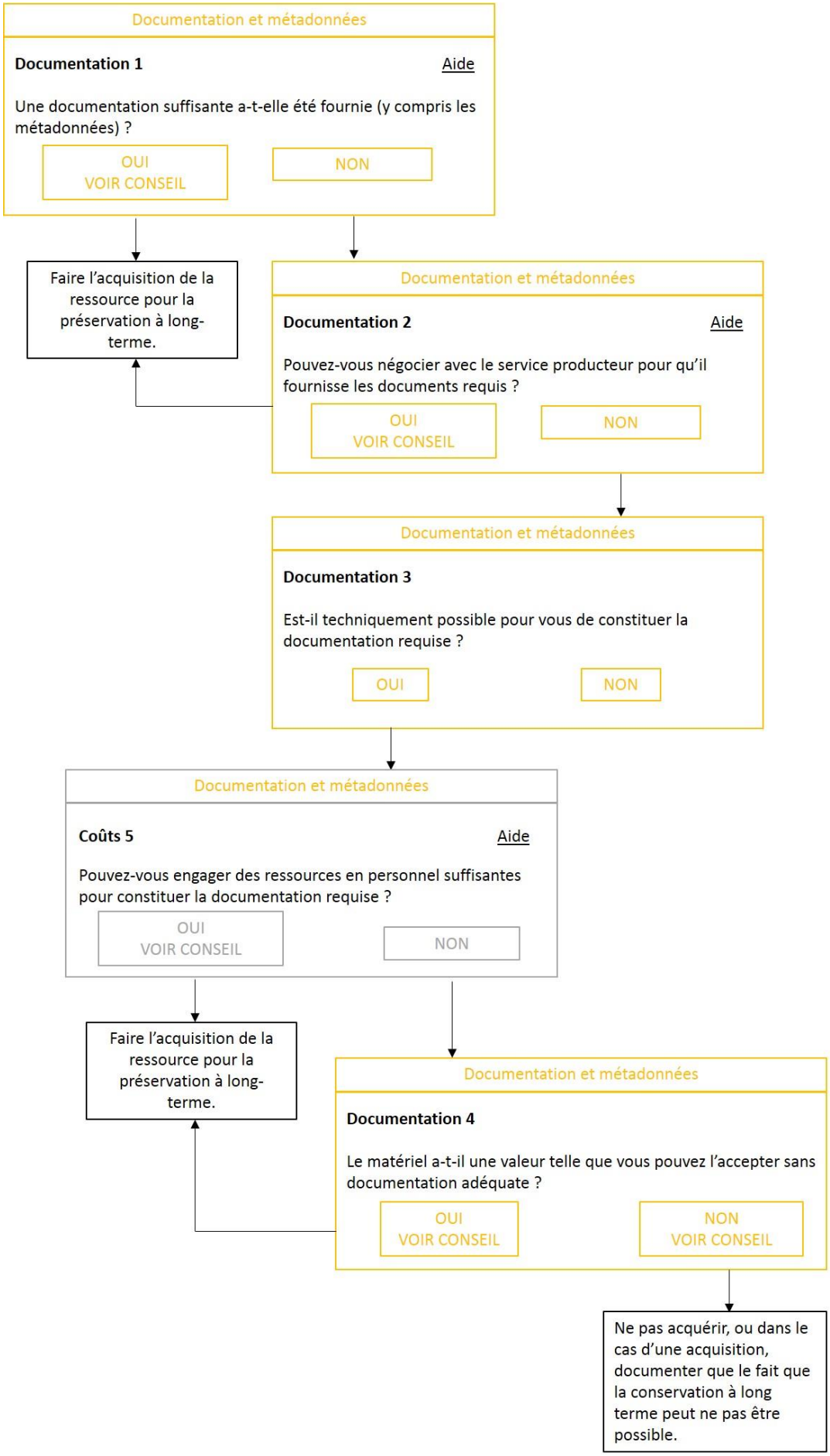
Évaluation interactive : sélection d'objets numériques pour une préservation à long terme


Choix de contenu et de version
Sélection 1
Aide
Le contenu de cette ressource s'inscrit-il dans le cadre de la politique de développement de la collection/des attributions de l'institution ?
OUI
NON
Choix de contenu et de version
Sélection 2
Existe-t-il une organisation plus appropriée pour préserver l'objet ?
OUI
NON
Choix de contenu et de version
Sélection 4
La ressource existe-t-elle sous plusieurs formes (par exemple numérique et papier) ?
OUI
NON POURSUIVRE A LA SECTION SUIVANTE
Choix de contenu et de version
Sélection 3
Avez-vous besoin de mener des acquisitions à d'autres fins, par exemple pour les informations que contient l'objet ?
OUI VOIR CONSEIL
NON VOIR CONSEIL
Mener l'acquisition pour utiliser et non préserver l'objet.
Ne pas réaliser l'acquisition.
Choix de contenu et de version
Sélection 5
Aide
La version numérique doit-elle être sélectionnée pour la préservation ?
OUI
NON VOIR CONSEIL
Mener l'acquisition pour utiliser et non préserver l'objet.
Choix de contenu et de version
Sélection 6
Aide
L'organisation peut-elle répondre aux exigences liées à l'éventuelle sensibilité des données ?
OUI POURSUIVRE A LA SECTION SUIVANTE
NON VOIR CONSEIL
Voir les parties « Sécurité de l'information » et « Respect de la législation »


Droit et responsabilités
Droit 1
L'acquisition confère-t-elle automatiquement une responsabilité en matière de conservation (par exemple, acquisition par dépôt légal/exigences en matière de confidentialité) ?
OUI
NON
Droit et responsabilités
Droit 2
La responsabilité de la préservation a-t-elle été prise en charge ailleurs ?
OUI
NON / NE SAIT PAS
Droit et responsabilités
Droit 3
Aide
Avez-vous besoin d'une garantie de préservation ou d'un accès à cette ressource plus important que celui offert ailleurs ?
OUI
NON
VOIR CONSEIL
Mener l'acquisition pour utiliser et non préserver l'objet.
Droit et responsabilités
Droit 4
Aide
Avez-vous (ou pouvez-vous négocier) des conditions acceptables pour l'acquisition et/ou le transfert, y compris la fréquence d'entrée appropriée à votre modèle d'organisation, ainsi que pour les modalités d'accès ?
OUI
NON
Droit et responsabilités
Droit 5
Réévaluer l'acquisition. La ressource vaut-elle la peine d'être acquise et préservée même s'il existe des restrictions indésirables ?
OUI
NON
VOIR CONSEIL
Ne pas acquérir, ou dans le cas d'une acquisition, documenter que le fait que la conservation à long terme peut ne pas être possible.
Droit et responsabilités
Droit 6
Avez-vous (ou pouvez-vous négocier) les droits de transfert sur différents supports et/ou différents formats à des fins de conservation ?
OUI
POURSUIVRE A LA SECTION SUIVANTE
NON
VOIR CONSEIL
Acquérir la ressource, mais documenter le fait que la conservation à long terme peut ne pas être possible.


Technique / Coûts
Technique 1
Aide
La ressource est-elle dans un format de fichier que vous pouvez gérer actuellement et dans l'avenir, ou pouvez-vous négocier pour que vous soit fournie ressource dans un format de fichier gérable ?
OUI
NON
Technique / Coûts
Technique 2
Est-il techniquement possible pour vous de migrer l'objet vers un format de fichier plus facile à gérer ?
OUI
NON
Technique / Coûts
Coûts 1
Aide
Êtes-vous prêt et capable d'accepter les coûts et les risques liés à la gestion de la ressource dans ce format de fichier ?
OUI
NON
VOIR CONSEIL
Acquérir la ressource, mais documenter le fait que la conservation à long terme peut ne pas être possible.
Technique / Coûts
Technique 3
La ressource est-elle disponible en ligne ou sur un support physique ?
EN LIGNE
PHYSIQUE
Voir page suivante
Technique / Coûts
Technique 4
Pouvez-vous collecter ou recevoir la ressource par une méthode appropriée et sûre (c'est-à-dire acquérir une copie de la ressource et pas seulement les droits d'accès à une copie) ?
OUI
NON
VOIR CONSEIL
Ne pas acquérir, ou dans le cas d'une acquisition, documenter que le fait que la conservation à long terme peut ne pas être possible.
Technique / Coûts
Technique 5
Pouvez-vous stocker et gérer l'objet dans des systèmes de stockage résilients à copies multiples ?
OUI
POURSUIVRE A LA
SECTION SUIVANTE
NON
Voir page suivante.


Technique / Coûts
Technique 5
Pouvez-vous stocker et gérer l'objet dans des systèmes de stockage résilients à copies multiples ?
OUI POURSUIVRE A LA SECTION SUIVANTE
NON
Technique / Coûts
Technique 3
La ressource est-elle disponible en ligne ou sur un support physique ?
EN LIGNE
PHYSIQUE
Voir page précédente.
Technique / Coûts
Coûts 2
Aide
Est-il rentable pour vous de développer un système de stockage résilient avec des copies multiples ?
OUI POURSUIVRE A LA SECTION SUIVANTE
NON VOIR CONSEIL
Ne pas acquérir, ou dans le cas d'une acquisition, documenter que le fait que la conservation à long terme peut ne pas être possible.
Technique / Coûts
Technique 6
Aide
La ressource se trouve-t-elle sur un support acceptable pour le transfert et/ou le stockage, ou pouvez-vous négocier pour que la source fournisse un support acceptable ?
OUI POURSUIVRE A LA SECTION SUIVANTE
NON
Technique / Coûts
Technique 7
Est-il techniquement possible pour vous de transférer la ressource vers un support acceptable ?
OUI
NON
Technique / Coûts
Coûts 3
Est-il rentable pour vous de transférer la ressource vers un support acceptable ?
OUI POURSUIVRE A LA SECTION SUIVANTE
NON
Technique / Coûts
Coûts 4
Aide
Êtes-vous prêt et capable d'accepter les coûts et les risques liés à la gestion de la ressource sur ce support ?
OUI POURSUIVRE A LA SECTION SUIVANTE
NON VOIR CONSEIL
Ne pas acquérir, ou dans le cas d'une acquisition, documenter que le fait que la conservation à long terme peut ne pas être possible.


Documentation et métadonnées
Documentation 1
Aide
Une documentation suffisante a-t-elle été fournie (y compris les métadonnées) ?
OUI
VOIR CONSEIL
NON
Faire l'acquisition de la ressource pour la préservation à long-terme.
Documentation et métadonnées
Documentation 2
Aide
Pouvez-vous négocier avec le service producteur pour qu'il fournisse les documents requis ?
OUI
VOIR CONSEIL
NON
Documentation et métadonnées
Documentation 3
Est-il techniquement possible pour vous de constituer la documentation requise ?
OUI
NON
Documentation et métadonnées
Coûts 5
Aide
Pouvez-vous engager des ressources en personnel suffisantes pour constituer la documentation requise ?
OUI
VOIR CONSEIL
NON
Faire l'acquisition de la ressource pour la préservation à long-terme.
Documentation et métadonnées
Documentation 4
Le matériel a-t-il une valeur telle que vous pouvez l'accepter sans documentation adéquate ?
OUI
VOIR CONSEIL
NON
VOIR CONSEIL
Ne pas acquérir, ou dans le cas d'une acquisition, documenter que le fait que la conservation à long terme peut ne pas être possible.

Liste des rubriques d'aide

Sélection 1
Si vous n'avez pas encore élaboré une politique de sélection des objets numériques pour la préservation à long terme, consultez les sections Politiques et stratégies institutionnelles et Acquisition et évaluation du manuel pour obtenir des conseils.
Les décisions relatives à la valeur à long terme prendront en compte les éléments suivants :

- Y a-t-il une exigence légale associée à l'objet ?
- L'objet a-t-il une grande valeur pour l'activité ou la mission de votre institution ?
- Le contenu de l'information est-il de bonne qualité ?
- Avez-vous ou pouvez-vous négocier des modalités d'accès acceptables ?
- Peut-on raisonnablement considérer que l'objet est un important représentant de sa typologie ?
- Est-il susceptible d'être utilisé pour de futures recherches ?

Sélection 5
Les décisions tiendront compte des facteurs suivants :

- Prévision du format préféré par les utilisateurs ;
- Considérations techniques concernant la facilité d'utilisation et la capacité de maintenance dans le temps ;
- Nécessité d'avoir la ressource à disposition à long terme ou uniquement à court terme ;
- Existence de variations significatives entre les formats pour qu'ils soient considérés comme des versions différentes ;
- Si l'une des versions offre de plus grandes possibilités d'utilisation.

Sélection 6
Voir les sections Sécurité de l'information et Respect de la législation.

Droit 3
Si la responsabilité de la préservation a été acceptée ailleurs, vous devez prendre en compte les éléments suivants pour décider ou non d'assumer également la responsabilité de la préservation :

- La responsabilité de la préservation a-t-elle été assumée uniquement à court terme ?
- L'infrastructure de l'autre organisation est-elle manifestement capable de garantir l'accès à long terme à la ressource ?
- Y a-t-il des conditions d'accès imposées que votre organisation jugerait inacceptables ?
- La responsabilité a-t-elle été assumée ailleurs uniquement pour diffuser la ressource, c'est-à-dire pour fournir un accès mais pas nécessairement pour garantir la préservation à long terme ?
- Quelles sont les approches de préservation (capture, stockage, etc.) utilisées ?

Droit
Voir la section Respect de la législation.

Technique 1
Voir les sections Anciens supports et Formats de fichiers et normes.

Technique 6
Voir les sections Anciens supports et Stockage.

Coûts 1
Voir les sections Anciens supports et Analyse de rentabilité, avantages, coûts et impact.

Coûts 2
Voir les sections Stockage et Analyse de rentabilité, avantages, coûts et impact.

Coûts 4
L'évaluation du rapport coût-bénéfice du reformatage des supports de stockage standard varie d'un établissement à l'autre. Les points suivants indiquent les facteurs qui peuvent être pris en compte avant de décider si cela est rentable :

- Stabilité du format ;
- Capacité de stockage des médias ;
- Volumétrie de l'objet ;
- Facilité d'accès si l'on prévoit une utilisation fréquente.
- Droit d'auteur. Si vous disposez de droits d'exception en matière d'archivage, le propriétaire vous a déjà autorisé à reformater vos documents pour les stocker dans d'autres supports en ligne ou hors ligne.
- Considérations opérationnelles (par exemple, il peut être plus efficace de tout reformater sur un support commun plutôt que de sélectionner seulement certaines catégories de données).

Coûts 5
Voir les sections Métadonnées et documentation et Analyse de rentabilité, avantages, coûts et impact.

Documentation 1
Voir la section Métadonnées et documentation

Documentation 2
Envisagez la rédaction de guides. Voir la section Métadonnées et documentation

Préservation et (ré)évaluation

Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

C'est normalement au moment de l'acquisition que l'on sélectionne ce qui doit être conservé sur le long terme. Mais la sélection peut également être un processus itératif, intervenant ultérieurement, une fois que les objets numériques se trouvent déjà conservés dans les collections. L'expression « préservation et (ré) évaluation » est ici utilisée pour désigner ce processus itératif. Le processus de décision reprend des mêmes étapes que celles décrites pour le processus d'acquisition et d'évaluation. L'arbre de décision présenté ci-dessus peut être adapté à cette fin.

Utiliser des critères d'évaluation et des procédures de sélection pour toutes les acquisitions potentielles d'objets numériques garantit que l'accroissement des collections est soigneusement priorisé et durable. Cela devrait réduire au minimum la fréquence et la nécessité des décisions de préservation et de réévaluation, car les acquisitions sont soigneusement évaluées et justifiées avant l'entrée dans les collections. Les organisations peuvent également avoir besoin de conserver certains documents internes et objets numériques pour des raisons réglementaires, juridiques, opérationnelles et financières. Ceux-ci doivent être gérés de manière active pour préserver leur viabilité, leur authenticité et leur accessibilité.

La préservation des objets numériques collectés au fil du temps et avant la mise en place de politiques et procédures institutionnelles devra normalement être réévaluée. Ce peut être l'une des premières étapes qu'une institution entreprend lors de la mise en œuvre d'une politique de préservation numérique : quantifier ses fonds numériques actuels et évaluer les risques de préservation (voir Pour commencer).

Les services d'archives utilisent le concept de série pour désigner un ensemble de documents qui partagent des caractéristiques similaires. Généralement, de nombreuses séries sont ouvertes sur des décennies. Cependant, le périmètre et la portée d'une série numérique peuvent changer au fil du temps et il est certain que les considérations technologiques sont susceptibles d'évoluer et qu'une certaine attention doit être accordée à une évaluation minutieuse lors de chaque transfert vers le service d'archives.

Avec le temps, il peut également s'avérer nécessaire de réévaluer les collections et la politique d'acquisition afin de refléter l'évolution des besoins et des circonstances. Dès lors qu'il est nécessaire de prendre des décisions précoces en matière de sélection de ce qui doit être préservé dans un environnement numérique (sans le recul qui est souvent disponible dans l'environnement analogique), il peut s'avérer nécessaire de procéder à une réévaluation dans le cycle de vie de préservation des objets numériques.

Dans une bibliothèque numérique où des niveaux de collection ont été définis, les objets numériques correspondant à chaque niveau peuvent être soumis à une réévaluation périodique, passés d'un niveau à l'autre, être retirés ou faire l'objet d'un désherbage en fonction de l'évolution des besoins et des circonstances. Cependant, s'agissant des objets sélectionnés pour une préservation permanente, la réévaluation et le désherbage interviennent rarement et de manière contrôlée. S'agissant des ressources redondantes ou sous licence, la réévaluation peut s'appuyer sur les critères suivants :

- Le contenu est désormais utilisé à un degré moindre que le niveau acceptable
- Le contenu est disponible ailleurs à un degré de qualité supérieur ou à un coût considérablement inférieur.

Les contenus qui ont été révisés ou qui ne sont plus suffisamment pertinents pour justifier un maintien sous forme active peuvent être conservés avec leurs mises à jour ultérieures ou retirés.

- Expiration ou résiliation d'une licence ou d'un accord d'échange de données ; retrait/retour d'une ressource numérique au fournisseur de données.
- Coût nécessaire pour maintenir la ressource dépasse la valeur/le bénéfice obtenu.
- Détérioration de la qualité du service fourni par un fournisseur ou détérioration de l'accessibilité du contenu en raison d'une mauvaise mise à jour de l'indexation, de l'imagerie ou d'autres caractéristiques internes à la ressource.

Pour les records managers et les archivistes, l'utilisation des durées et des calendriers de préservation est bien établie. Les documents d'archives peuvent être détruits à la fin de leur période de préservation, conservés pendant une période supplémentaire ou transférés à une autre institution en vue d'une préservation à long terme.

Dans tout organisme de préservation, il est important que des procédures écrites soient définies pour le processus de préservation et de réévaluation, avec des responsabilités clairement attribuées à des personnes désignées ou aux sections d'une organisation chargées de la gouvernance et du développement des collections. Les délais, les circonstances et les procédures d'autorisation pour la réévaluation doivent être clairement indiqués. Les calendriers de préservation et de réévaluation eux-mêmes doivent être périodiquement révisés pour suivre le rythme des nouvelles exigences organisationnelles. Selon le contexte de fonctionnement de l'institution, ses utilisateurs et déposants peuvent être consultés dans le cadre de ce processus. Tout avis peut ensuite être soumis à l'approbation de la direction et de comités, en fonction de la taille et de l'importance de la ressource.

Lorsqu'un avis est donné de retirer une ressource archivée, doivent exister des procédures permettant de consulter les autres parties prenantes afin de déterminer si un transfert vers une autre organisation s'avère pertinent. Dans ce cas, l'institution doit convenir des conditions de transfert, y compris les niveaux acceptables d'attention portée à la ressource et l'accès fourni à celle-ci, selon les besoins des utilisateurs du monde de l'enseignement et de la recherche. Les contraintes financières ne devraient pas être le principal moteur de la sortie définitive ou du désherbage des objets numériques, car le processus lui-même n'est pas sans coût et peut soulever

d'épineuses questions éthiques. Les décisions de sortie définitive et de désherbage devraient être principalement motivées par la politique de développement des collections.
Les objets numériques collectés qui n'ont pas été conservés après réévaluation doivent être inscrits dans un catalogue institutionnel avec des commentaires indiquant le processus mis en œuvre et les détails du transfert.

Planifier le transfert à une autre institution peut également s'avérer pertinent dans le cadre de la préservation et de la réévaluation. La recommandation 3.1.2.1 de la norme pour l'audit et la certification des entrepôts numériques dignes de confiance (CCSDS, 2011) exige que l'entrepôt numérique dispose d'un plan de transfert approprié, de plans d'urgence et/ou d'accords de séquestre au cas où l'entrepôt numérique cesserait de fonctionner ou que l'institution directrice ou de financement en modifierait considérablement la portée.

Ressources

Deaccessioning and disposal. Guidance for archive services.
http://www.nationalarchives.gov.uk/documents/Deaccessioning-and-disposal-guide.pdf
Ce guide a été rédigé pour accompagner l'*Archive Service Accreditation*, un standard britannique pour les services d'archives. Il est générique et s'applique au contenu dans tous les formats, analogique comme numérique, et ne comporte aucune disposition particulière pour les documents numériques. La norme présente la sortie définitive comme faisant partie du développement des collections et exige que les services d'archives aient des politiques, des plans et des procédures en place pour les activités de développement des collections, y compris concernant la sortie définitive. Il comprend un arbre de décision pour identifier le sort final (2015, 34 pages).

Digital Content Review: Process and Results
http://dpworkshop.org/workshops/management-tools/process-results
Pour réévaluer des objets numériques, l'équipe de préservation numérique rassemble des informations et les accumule de manière itérative dans le cadre d'un processus structuré. Les résultats des opérations de réévaluation des objets numériques en cours produisent un ensemble de données de réévaluation des objets numériques. Ces données permettent aux organisations de planifier leurs opérations à court et à long terme.

Disposing of Records
http://www.nationalarchives.gov.uk/information-management/manage-information/policy-process/disposal/
Les Archives nationales du Royaume-Uni ont produit un guide étape par étape pour vous aider dans le processus de mise en œuvre du sort final, y compris une liste de contrôle.

Deaccessioning policy
http://www.bl.uk/aboutus/stratpolprog/coldevpol/deaccessioning/
Cette politique définit les circonstances dans lesquelles la *British Library* peut procéder au désherbage de certains types d'objets. Il s'agit d'une politique générique qui s'applique à des contenus de tous formats, analogique ou numérique, et qui ne comporte pas de dispositions particulières pour les objets numériques.

Références

Consultative Committee for Space Data Systems (CCSDS), 2011.*Audit and Certification of Trustworthy Digital Repositories Recommended Practices CCSDS 652.0-M-1 Magenta Book September 2011* [publié ensuite sous la forme de la norme ISO 16363: 2012]. Disponible à l'adresse suivante : https://public.ccsds.org/pubs/652x0m1.pdf

Stockage

Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Introduction

Cette section traite de l'émergence de l'utilisation des systèmes de stockage informatique pour la préservation numérique. Elle traite de questions génériques, sachant que les questions plus spécifiques associées au stockage dans le *cloud* sont traitées séparément (voir la section Services *cloud*). La pratique traditionnelle de préservation des supports numériques, par exemple les anciens supports des collections existantes, est également abordée ailleurs (voir la section Anciens supports). De nombreuses organisations ont des stratégies mixtes ou sont en train de passer de l'une à l'autre.

L'utilisation des technologies de stockage pour la préservation numérique a considérablement changé au cours des vingt dernières années. Au cours de cette période, les pratiques ont changé. Auparavant, la norme consistait à stocker des objets numériques à l'aide d'éléments de support séparés, par exemple des CD individuels, des bandes, etc. qui étaient ensuite migrés périodiquement pour remédier à la dégradation et à l'obsolescence. Aujourd'hui, il est devenu plus courant d'utiliser des systèmes de stockage informatique résilients pour les volumes de plus en plus importants d'objets numériques à préserver et, peut-être plus important encore, ces systèmes doivent être facilement et rapidement récupérables à un moment où la pratique de l'accès en ligne s'est généralisée. De cette manière, les objets numériques ont été dissociés de leur mécanisme sous-jacent de son stockage, ce qui a l'avantage de permettre à différentes activités de préservation d'être réalisées indépendamment.

Systèmes de stockage résilients

Un système de stockage informatique résilient est constitué de supports de stockage sur un serveur qui offre une résilience à divers types de pannes en utilisant un système intégré de redondance et de récupération. Par exemple, un système de stockage peut être constitué de disques RAID (*Redundant Array of Independent Disks*), de bandes dans une bibliothèque, ou d'une combinaison de plusieurs types de stockage dans un système de gestion hiérarchique (HSM, *Hierarchical Storage Management*). Il peut s'agir d'un stockage sur site et/ou d'un stockage à distance sur le *cloud* et d'une réplication automatisée des objets numériques sur plusieurs sites et systèmes.

Ces systèmes deviendront toujours obsolètes au fil du temps et les objets numériques doivent être transférés régulièrement d'un système à un autre à mesure que ceux-ci deviennent obsolètes. La migration entre systèmes de stockage est distincte de la migration de formats de fichiers et peut être traitée en grande partie comme un problème informatique, à condition qu'une surveillance adéquate soit exercée pour garantir le respect des exigences de préservation. L'avantage est que l'utilisation de systèmes informatiques pour le stockage des données peut permettre un accès beaucoup plus rapide, une solution plus évolutive, une gestion plus facile et, en fin de compte, des coûts plus faibles, surtout en cas de mise à l'échelle.

Il est essentiel de comprendre la différence entre les solutions de stockage informatique standard et les besoins supplémentaires induits par la préservation à long terme. Il est essentiel de pouvoir expliquer ces différences à votre service informatique ou à votre fournisseur de services de stockage et de pouvoir spécifier ces exigences lors de l'acquisition d'un système ou d'un service. Les systèmes de stockage standard sont conçus pour des objets numériques en cours d'utilisation. Bien que des procédures de sauvegarde soient généralement prévues, elles ne répondent généralement pas aux exigences plus strictes visant à garantir la préservation à long terme des objets numériques. Sauvegarde et préservation numérique ne sont pas la même chose et de nombreux services ou experts informatiques peuvent ne pas comprendre ce point. Les systèmes de stockage pour la préservation numérique exigent un niveau plus élevé de redondance géographique, des procédures de reprise après sinistre renforcées, une planification à plus long terme et, surtout, une surveillance active de l'intégrité des données afin de détecter les modifications indésirables telles que la corruption ou la perte de fichiers.

Il existe de nombreuses façons de répondre aux besoins de stockage en matière de préservation et celles-ci varient en termes d'échelle et de complexité en fonction du contexte organisationnel. Il sera nécessaire d'évaluer les ressources internes et d'envisager des options d'externalisation et de stockage dans le *cloud*. L'approche adoptée dépendra souvent de la taille et de la complexité de la collection, ainsi que des ressources disponibles au sein de l'organisation. Il est possible de répondre aux besoins de stockage pour la préservation avec une configuration de base, mais à mesure que la taille d'une collection augmente, il sera nécessaire d'aborder des questions telles que la scalabilité et l'automatisation.

Principes d'utilisation des systèmes de stockage informatique pour la préservation numérique	
Les éléments suivants représentent les principes qui devraient être utilisés lors de la conception ou du choix des systèmes de stockage pour la préservation.	
1	**Redondance et diversité** • Faire plusieurs copies indépendantes des objets numériques et les stocker dans différents lieux géographiques. • Utiliser une combinaison de systèmes de stockage en ligne et de supports hors ligne. • Utiliser différentes technologies de stockage pour répartir les risques et trouver un équilibre entre la sécurité des données, la facilité d'accès et un coût raisonnable.
2	**Intégrité, surveillance, réparation** • Utiliser des mécanismes comme les empreintes pour enregistrer et contrôler régulièrement l'intégrité de chaque copie d'un objet numérique. • Si une corruption ou une perte est détectée, utilisez l'une des autres copies pour remplacer la copie défaillante ou perdue. • Stocker les informations relatives à l'intégrité avec les objets numériques et également dans des bases de données ou des systèmes séparés.
3	**Surveillance des technologies et des fournisseurs, évaluation des risques et migrations proactives** • Comprendre que les technologies, produits et services de stockage ont tous une courte durée de vie. • Utiliser la veille technologique pour évaluer le moment où des migrations pourraient être nécessaires. • Garder un œil sur la viabilité des fournisseurs de stockage ou des catégories de solutions de stockage. • Être proactif en matière de migration du stockage avant que le matériel numérique ne soit considéré à risque.
4	**Consolidation, simplification, documentation, provenance et suivi d'audit** • Réduire au minimum la prolifération des types de supports anciens et des systèmes de stockage utilisés pour la préservation. • Regrouper les objets numériques sur le minimum de systèmes de stockage (sous réserve des exigences de redondance ci-dessus). • Documenter la manière dont les objets numériques ont été collectés et transférés sur les systèmes de stockage ainsi que la manière dont les systèmes de stockage sont mis en place et exploités. • Utiliser cette documentation pour fournir des informations permettant d'auditer l'authenticité des données.

Fiabilité du stockage

Lorsque l'on examine les solutions de stockage, sur site ou dans le *cloud*, on se demande dans quelle mesure elles sont fiables et si quelque chose ne va pas, ce que cela signifie en termes de perte de données. Les fabricants font généralement valoir des statistiques en matière de fiabilité, de durabilité, de taux d'échec et de taux d'erreur.

Par exemple, cela peut prendre la forme d'un service *cloud* conçu pour une durabilité de 99,999 %, d'un taux d'erreur (*Bite Error Rate* - BER) de 1 sur 10^{16} lors de la lecture des données sur bande, ou d'un temps moyen entre deux pannes (MTBF) de 1 million d'heures pour un disque dur. Ces chiffres sont ensuite utilisés pour calculer d'autres mesures de fiabilité. Par exemple, un temps moyen avant perte de données (MTTDL) de 1 000 ans peut être mis en avant lorsque des disques durs sont utilisés dans une matrice RAID6.

Ces chiffres sont souvent difficiles à comprendre et doivent être interprétés avec beaucoup de prudence lorsqu'on tente d'estimer le degré de sécurité d'une solution de stockage donnée.
Des travaux importants ont déjà été réalisés sur la manière de décrire, de mesurer et de prévoir la fiabilité du stockage, notamment du point de vue de la préservation numérique. Il s'agit d'un sujet complexe qu'il n'est pas possible d'aborder dans ce manuel. Pour en savoir plus, voir par exemple Greenan et al (2010), Rosenthal (2010) et Elerath (2009). Plusieurs considérations importantes ressortent de ces travaux :

- Les technologies de stockage informatique sont en général remarquablement fiables pour ce qu'elles font. Les défaillances sont relativement rares, mais elles se produisent et se produiront. Il est tentant de supposer que ce n'est pas parce qu'un type particulier de panne n'a pas été rencontré dans un cas que la technologie de stockage est en général plus fiable qu'elle ne l'est réellement. C'est un point de vue dangereux. Par exemple, de nombreuses personnes possèdent des disques durs qui fonctionnent parfaitement bien depuis des années et des années, mais la réalité est qu'en moyenne, jusqu'à 5 à 15 % des disques durs tombent en panne dans l'année qui suit (Backblaze, 2014), (Pinheiro et al, 2007).
- Les défaillances et les erreurs étant des événements relativement rares, les statistiques de fiabilité des fournisseurs sont généralement basées sur des modèles et des simulations et non sur des observations à long terme de ce qui se passe réellement dans la pratique. Par exemple, si un fabricant déclare que la durée de vie d'un support est de 30 ans, ce n'est pas parce qu'il a réellement testé le support pendant cette période. De même, si un fournisseur estime la durée de vie d'un support à 1 000 ans, c'est qu'il n'a manifestement pas construit un système et ne l'a pas testé pendant une période aussi longue. Par conséquent, les statistiques doivent être interprétées comme les meilleures estimations des fournisseurs sur la façon dont un système pourrait se comporter dans l–pratique - mais il se peut qu'il n'en soit pas ainsi. Par exemple, des études de terrain ont suggéré que les estimations de fiabilité des fabricants peuvent être trop optimistes (Jiang et al., 2008).
- La probabilité de perte de données augmente considérablement lorsque les corrélations sont prises en compte. On parle de corrélations lorsque des parties d'un système, ou différentes copies d'un objet numérique, ne peuvent pas être considérées comme indépendantes. Si un problème survient sur une partie du système ou une copie de l'objet numérique, il est probable qu'il y ait un problème avec une autre partie ou une autre copie. Il peut s'agir par exemple d'un défaut de fabrication affectant tous les disques durs d'un serveur de stockage, de bugs logiciels ou micrologiciels corrompant systématiquement l'objet numérique, de l'incapacité d'une organisation à tester régulièrement ses sauvegardes, ou de l'incapacité à isoler ou découpler les systèmes de stockage de sorte que si une copie de l'objet numérique est accidentellement supprimée, toutes les autres copies ne le soient pas non plus. Ces facteurs de corrélation peuvent être beaucoup plus graves que les modes de défaillance spécifiques couverts par les statistiques de fiabilité.

Ces constats et observations conduisent à faire les recommandations suivantes :

- Prévoir les défaillances des solutions de stockage informatique, quelle que soit l'ingéniosité du fabricant. En pratique, les taux de défaillance pourraient bien être plus élevés que ce que les statistiques des fabricants laissent entendre.
- La perte de données peut être causée par l'incapacité à mettre en place des processus et des procédures appropriés autour de l'utilisation du stockage informatique ainsi que par la technologie de stockage elle-même. Une évaluation appropriée des risques est le meilleur moyen d'identifier et de gérer ces problèmes.
- La meilleure stratégie consiste à créer plusieurs copies indépendantes des objets numériques dans différents endroits et de les stocker en utilisant différentes technologies lorsque cela est possible. Cela devrait inclure un processus de contrôle actif et régulier de l'intégrité des données de toutes les copies afin que les problèmes puissent être détectés quels que soient la raison et le lieu où ils se produisent. De cette façon, les risques sont à la fois minimisés et répartis, et on ne peut pas compter sur une technologie ou un service de stockage particulier qui soit totalement exempt d'erreurs.

Stratégies de stockage de copies multiples

Les technologies de stockage numérique présentent plusieurs risques pour la préservation à long terme des objets numériques. Ces risques peuvent être réduits en utilisant une stratégie de stockage numérique qui implique l'utilisation d'un ou de plusieurs systèmes de stockage et au moins de deux copies des données.

La bonne pratique veut qu'une stratégie de stockage présente les caractéristiques suivantes :

a) il existe de multiples copies indépendantes des objets numériques
b) ces copies sont séparées géographiquement en différents endroits
c) les copies utilisent des technologies de stockage différentes
d) les copies utilisent une combinaison de techniques de stockage en ligne et hors ligne
e) le stockage est activement surveillé pour garantir que tout problème est détecté et corrigé rapidement.

Une stratégie de stockage numérique peut être mise en œuvre de manière progressive, en commençant par un niveau de protection et d'accès basiques au contenu numérique, puis en passant à une approche plus automatisée et scalable qui offre un niveau plus élevé de sécurité et de sûreté des données.

Les risques liés aux contenus numériques proviennent de diverses sources et une stratégie de stockage numérique permet de trouver un équilibre entre le coût du stockage numérique et la réduction de ces risques. Parmi les exemples de risques à prendre en compte figurent les incendies, les inondations, l'incapacité à mettre en place ou à suivre des processus ou des procédures appropriées, les attaques malveillantes, la dégradation des supports et l'obsolescence des systèmes et des technologies de stockage. Les principaux risques et les moyens de les traiter ou de les atténuer sont souvent abordés dans le plan de continuité des activités d'une organisation (voir Gestion des risques et du changement).

Il est important de réaliser que de nombreux exemples de perte de contenu ne sont pas nécessairement dus à des défaillances techniques de la technologie de stockage (bien qu'il soit important de reconnaître que celles-ci se produisent), mais peuvent provenir d'une erreur humaine, d'un manque de budget ou de planification des migrations de stockage, ou d'un manque de vérification et de correction régulières des défaillances qui pourraient survenir.

Dans un monde qui utilise de plus en plus de systèmes et de technologies en réseau pour le stockage numérique, une copie hors ligne des objets numériques a un rôle à jouer. Cela peut constituer un « coupe-feu » contre les problèmes des systèmes en ligne qui peuvent se propager automatiquement d'un endroit à l'autre, par exemple la suppression d'un fichier à un endroit qui supprime automatiquement une copie en miroir sur un autre site.

Il est essentiel de faire plusieurs copies des objets numériques pour atteindre un niveau basique de sécurité des données. L'utilisation de différents types de stockage pour chaque copie permet de répartir le risque et de s'assurer qu'un problème avec une technologie n'affecte pas les autres. La manière dont chaque copie est stockée peut être ajustée pour atteindre un niveau global acceptable de coût, de risque et de complexité. Par exemple, une copie peut être conservée sur un serveur de stockage en ligne pour un accès rapide et une copie peut être stockée sur bande en stockage froid pour un faible coût et un niveau de sécurité relativement élevé.

Ce manuel suit les niveaux de préservation de la *National Digital Stewardship Alliance* (NDSA) (NDSA, 2013) en recommandant ci-dessous quatre niveaux auxquels la préservation numérique peut être mise en œuvre via le stockage et la redondance géographique. Nous recommandons en outre d'utiliser une combinaison de copies en ligne et hors ligne pour parvenir à un bon compromis entre accès et sécurité des données :

Niveau	Approche	Risques traités et avantages obtenus
1	• Deux copies complètes des objets numériques non co-localisées. Une copie doit être hors ligne. • Pour les objets numériques sur des supports hétérogènes (disques optiques, disques durs, etc.), extraire le contenu du support et l'écrire dans votre système de stockage.	• Capacité basique à se remettre d'une série de problèmes, notamment d'une défaillance du système de stockage. La perte ou l'endommagement d'une copie peut être récupéré en utilisant l'autre copie. • Les objets numériques sont plus faciles à gérer lorsqu'ils se trouvent dans un seul système de stockage
2	• Au moins trois copies complètes. • Au moins une copie sur un site géographique différent. • Documenter votre (vos) système(s) de stockage et vos supports de stockage et ce dont vous avez besoin pour les utiliser.	• Comme ci-dessus, plus la protection contre les catastrophes naturelles et autres événements majeurs. • Bon niveau d'accès et de sécurité des objets numériques. • Le personnel dispose de politiques et de procédures claires à suivre pour être plus efficace, les coûts sont réduits et les changements de personnel peuvent être gérés.
3	• Au moins une copie sur un site géographique avec une menace de catastrophe différente. • Surveiller l'obsolescence et disposer d'un processus de migration pour votre (vos) système(s) de stockage et vos supports.	• Comme ci-dessus, plus une protection contre les risques à long terme liés à l'obsolescence technique. • L'accès continu aux contenus est possible même pendant les migrations et les catastrophes.
4	• Au moins trois copies sur des sites géographiques présentant des	• Comme ci-dessus, toute la gamme des risques est abordée, y compris les

	menaces de catastrophes différentes. • Disposer d'un plan global qui permettra de conserver les fichiers et les métadonnées sur les supports ou systèmes actuellement accessibles.	pertes accidentelles et les attaques malveillantes, le verrouillage des fournisseurs et les instabilités budgétaires. • Le contenu est hautement disponible, les coûts sont prévisibles et gérables, il est possible d'obtenir la certification en tant qu'entrepôt digne de confiance

Gérer l'obsolescence des systèmes de stockage et les risques

L'utilisation des technologies et des solutions de stockage nécessite une planification et une gestion minutieuses afin de disposer d'une approche efficace permettant de maintenir la préservation numérique. Si elle est bien menée, elle peut se traduire par de très bons niveaux de sécurité des données, un accès rapide aux contenus en cas de besoin et des coûts à la fois faibles et prévisibles.

Les technologies de stockage informatique peuvent tomber en panne ou provoquer la corruption des données et la durée de vie des supports et des systèmes est le plus souvent courte, en général 3 à 5 ans, ce qui signifie que les solutions deviennent rapidement obsolètes et qu'une migration est nécessaire pour éviter que les objets numériques ne soient menacés. Dans ce contexte, la migration consiste à déplacer les données d'un ancien système de stockage vers un nouveau système de stockage. Les objets numériques eux-mêmes ne changent pas, mais la solution de stockage, si. Un service informatique ou un fournisseur de services de stockage pensera à la migration au niveau du stockage. Par contre, dans le cas d'une migration de format de fichier, le format de fichier change, mais la manière dont les fichiers sont stockés ne change pas.

Ressources

NDSA Levels of Preservation

http://www.digitalpreservation.gov/ndsa/activities/levels.html (2013)
https://ndsa.org//publications/levels-of-digital-preservation/ (Version 2.0, 2018)

Les « niveaux de préservation numérique » de la *National Digital Stewardship Alliance* (NDSA) sont un ensemble de recommandations à plusieurs niveaux sur la manière dont les organisations devraient commencer à mettre en place ou à améliorer leurs activités de préservation numérique. Il s'agit d'un ensemble de directives relativement faciles à utiliser, utiles non seulement pour ceux qui commencent à penser à la préservation de leurs collections numériques, mais aussi pour les institutions qui planifient les prochaines améliorations à apporter à leurs systèmes et processus de travail de préservation numérique existants. Il n'est pas conçu pour évaluer la solidité des programmes de préservation numérique dans leur ensemble, car il ne couvre pas des éléments tels que les politiques, la dotation en personnel ou le soutien organisationnel.

Traduction française de la version 2.0 : https://hal-bnf.archives-ouvertes.fr/hal-02551807/document

On trouvera ci-dessous quelques-uns des plus notables fournisseurs de stockage informatique basiques, ainsi que des fournisseurs de services de préservation numérique spécialisés qui peuvent fournir un stockage sur site ou dans le *cloud* (voir également les services dans le *cloud*). Ces spécialistes peuvent généralement prendre en charge d'autres fonctions de préservation en plus du stockage.

Arkivum
http://arkivum.com

Digital Preservation Network
http://www.dpn.org

DSpace
http://www.dspace.org

ePrints
http://www.eprints.org

Fedora
http://fedorarepository.org

iRods
http://irods.org

LOCKSS
http://www.lockss.org

OCLC Digital Archive CONTENTdm
http://www.oclc.org/digital-archive.en.html

Portico
http://www.portico.org/digital-preservation/

Preservica
http://preservica.com

Rosetta
https://www.exlibrisgroup.com/products/rosetta-digital-asset-management-and-preservation/

Community Owned digital Preservation Tool Registry COPTR
http://coptr.digipres.org/Main_Page

Bien que principalement axé sur les outils, le registre COPTR couvre également une série de systèmes et de services de stockage. Il agit principalement comme un outil de recherche et d'évaluation permettant d'aider les professionnels à trouver les outils dont ils ont besoin pour préserver leurs données numériques. COPTR fournit des détails factuels basiques sur un outil, ce qu'il fait, comment trouver plus d'informations (URL pertinentes) et des références aux expériences des utilisateurs avec l'outil.

Le blog de DSHR
http://blog.dshr.org

David Rosenthal est informaticien et responsable scientifique du projet LOCKSS. Son blog couvre régulièrement les évolutions du stockage informatique et les tendances et implications pour la préservation numérique.

Études de cas

Case Study: Bodleian Library, University of Oxford
http://www.nationalarchives.gov.uk/documents/archives/case-study-oxford.pdf
Cette étude de cas porte sur la *Bodleian Library* et l'Université d'Oxford, et sur la mise en place d'une infrastructure locale « en *cloud* privé » pour ses collections numériques comprenant des livres, des images et du multimédia numérisés, des données de recherche et des catalogues. Elle présente le contexte organisationnel, la nature de ses exigences et approches en matière de préservation numérique, ses services de stockage, son infrastructure technique, ainsi que l'analyse de rentabilité et le financement. Elle se termine par les principaux enseignements tirés et les projets futurs. Janvier 2015 (4 pages).

Case Study: The Parliamentary Archives
http://www.nationalarchives.gov.uk/documents/archives/case-study-parliament.pdf
Cette étude de cas porte sur les Archives parlementaires. Il s'agit d'un exemple de service d'archives utilisant un ensemble hybride de solutions de stockage (en partie dans le *cloud* public et en partie installées localement) pour la préservation numérique, car le service d'archives dispose d'un système de préservation installé localement (*Preservica Enterprise Edition*) qui est intégré au stockage dans le *cloud* et au stockage local et stocke les objets sensibles localement, et non dans le *cloud*. Janvier 2015 (4 pages).

Case Study: Tate Gallery
http://www.nationalarchives.gov.uk/documents/archives/case-study-tate-gallery.pdf
Cette étude de cas traite de l'expérience de développement d'un entrepôt numérique partagé pour les quatre sites physiques de la Tate, basé sur le système de stockage commercial d'Arkivum. Elle présente le contexte organisationnel, la nature de leurs besoins et approches en matière de préservation numérique, et les raisons pour lesquelles ils ont choisi la solution sur site d'Arkivum, Arkivum/OnSite, de préférence à toute offre basée sur le *cloud*. Elle se termine par les principaux enseignements tirés et examine les plans de développement futur. Janvier 2015 (4 pages).

Références

Backblaze, 2014. Hard Drive Reliability Update – Sep 2014. *Backblaze*. [blog] Disponible à l'adresse suivante : https://www.backblaze.com/blog/hard-drive-reliability-update-september-2014/

Elerath, J., 2009. Hard-Disk Drives: The Good, the Bad, and the Ugly. *Communications of the ACM*. 52 (6), 38-45. Disponible à l'adresse suivante : doi:10.1145/1516046.1516059. http://cacm.acm.org/magazines/2009/6/28493-hard-disk-drives-the-good-the-bad-and-the-ugly/fulltext

Greenan, K.M., Plank, J.S. & Wylie, J.J., 2010.Mean time to meaningless: MTTDL, Markov models, and storage system reliability. *Proceedings of the 2nd USENIX conference on Hot topics in storage*

and file systems. Disponible à l'adresse suivante: https://www.usenix.org/legacy/event/hotstorage10/tech/full_papers/Greenan.pdf

Jiang, W. et al., 2008. *Are Disks the Dominant Contributor for Storage Failures? A Comprehensive Study of Storage Subsystem Failure Characteristics. Proceedings of the 6th USENIX Conference on File and Storage Technologies (FAST '08)*. Disponible à l'adresse suivante : http://www.usenix.org/events/fast08/tech/jiang.html

NDSA, 2013. T*he NDSA Levels of Digital Preservation: An Explanation and Uses, version 1 2013*. National Digital Stewardship Alliance. Disponible à l'adresse suivante : http://www.digitalpreservation.gov/ndsa/working_groups/documents/NDSA_Levels_Archiving_2013.pdf
Traduction française de la version 1 : https://hal-bnf.archives-ouvertes.fr/hal-02162334
Pinheiro, P., Weber, W-D.& Barroso, L.A., 2007. Failure Trends in a Large Disk Drive Population. *Proceedings of the 5th USENIX Conference on File and Storage Technologies (FAST' 07)*. Disponible à l'adresse suivante : http://static.googleusercontent.com/media/research.google.com/en//archive/disk_failures.pdf

Rosenthal, D.S.H., 2010. Bit Preservation: A Solved Problem? *The International Journal of Digital Curation*.5 (1) Stanford University Libraries, CA. Disponible à l'adresse suivante: http://www.ijdc.net/index.php/ijdc/article/view/151/224

Anciens supports

Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Introduction

De nombreuses organisations auront de grandes quantités de données stockées sur d'anciens supports, tels que les supports magnétiques et optiques, et des données continueront à être reçues sur d'anciens supports. En fin de compte, la meilleure stratégie à long terme pour la préservation des données sera la migration vers un stockage sous forme de fichiers et une gestion active par la suite (voir la section Stockage). Souvent, les supports originaux continuent à être préservés en parallèle, il est donc nécessaire de comprendre leurs exigences en matière de préservation et de stockage. Pour les organisations possédant d'importantes collections d'anciens supports, comprendre les risques auxquels chaque type de support est confronté facilitera également l'établissement de priorités pour la migration des collections et l'application d'outils et de méthodes de criminalistique numérique sera également utile (voir la section Criminalistique numérique).

Pour la préservation des supports magnétiques et optiques, deux aspects doivent être pris en compte : le support lui-même et le matériel et les logiciels nécessaires à son interprétation. Dans certains cas, le second aspect sera le plus difficile à traiter. À mesure que la popularité d'un support diminue, la fabrication de matériel cesse et celui-ci devient plus difficile à obtenir et à maintenir.

Préserver les anciens supports

Dans la plupart des cas, le moyen le plus simple de réduire les risques liés aux supports de stockage consiste à transférer tous les contenus sur un système de stockage administré. Cela signifie que le contenu peut être géré sans lien avec le support de stockage d'origine. C'est probablement suffisant pour la grande majorité des contenus numériques nécessitant une préservation. Toutefois, il peut y avoir quelques cas où il est nécessaire de conserver le support de stockage original d'une manière ou d'une autre. Dans certains cas, le support de stockage peut simplement être conservé comme un artefact, sans attente d'accès à long terme, par exemple lorsqu'il fait partie d'une collection hybride ou qu'il a une certaine valeur comme partie d'un ensemble (par exemple, une partie des collections d'un auteur éminent). Toutefois, lorsqu'un accès

continu au contenu est nécessaire, il convient de réfléchir soigneusement à la manière dont il pourrait être accessible à l'avenir.

Nous savons par expérience que les types de supports de stockage numérique changent fréquemment au fil du temps. Par exemple, la version précédente de ce manuel donnait un aperçu des supports de stockage magnétiques et optiques et fournissait des estimations de la durée de vie de certains types de supports qui étaient populaires au milieu des années 1990 (un manuel de préservation numérique rédigé au cours des décennies précédentes aurait probablement inclus des évaluations de cartes perforées et de bandes de papier). Compte tenu des tendances actuelles en matière de technologies de stockage, il est peut-être préférable de fournir maintenant un cadre qui permette l'évaluation continue des supports de stockage, qui pourraient désormais inclure les clés USB ou les disques durs externes. Un tel cadre a été fourni par les Archives nationales du Royaume-Uni (Brown, 2008). Il utilise une approche par tableaux de bord et indicateurs, évaluant les supports de stockage sélectionnés en fonction de six critères :

- Longévité (par exemple, durée de vie opérationnelle prouvée)
- Capacité
- Viabilité (par exemple, en termes de préservation de l'intégrité des preuves)
- Obsolescence
- Coût
- Sensibilité (par exemple, aux dommages physiques et aux différentes conditions environnementales).

Dans la pratique cependant, ce type d'évaluation n'est pas très abouti. Il existe de plus en plus de preuves qui suggèrent que la variation de la qualité de fabrication joue également un rôle majeur dans la longévité des supports (Harvey, 2011). C'est pourquoi, en fin de compte, la préservation numérique dépend normalement du transfert du contenu vers un système de stockage administré.

Ressources

Selecting Storage Media for Long-Term Preservation. Digital Preservation Guidance Note 2.
https://www.nationalarchives.gov.uk/documents/selecting-storage-media.pdf
Ce document fait partie d'une série de notes d'orientation produites par les Archives nationales du Royaume-Uni, donnant des conseils généraux sur les questions relatives à la préservation et à la gestion des documents d'archives électroniques. Il est destiné à être utilisé par toute personne impliquée dans la création de documents d'archives électroniques. Il fournit des informations aux producteurs et aux gestionnaires de documents d'archives électroniques sur la sélection des supports de stockage physique dans une perspective de préservation à long terme. Ces directives sont en vigueur depuis août 2008 (7 pages).

Care, Handling and Storage of Removable Media, Digital Preservation Guidance Note 3.
http://www.nationalarchives.gov.uk/documents/information-management/removable-media-care.pdf
Ce document fait partie d'une série de notes d'orientation produites par les Archives nationales du Royaume-Uni, donnant des conseils généraux sur les questions relatives à la préservation et à la gestion des documents d'archives électroniques. Il fournit des conseils sur le soin, la manipulation

et le stockage des supports de stockage amovibles. La note d'orientation est datée d'août 2008 (10 pages).

You've Got to Walk Before You Can Run: First Steps for Managing Born-Digital Content Received on Physical Media.
http://www.oclc.org/content/dam/research/publications/library/2012/2012-06.pdf
Un guide étape par étape pour extraire des objets numériques natifs de divers supports physiques. Il se concentre sur l'identification et la stabilisation de vos collections afin que vous soyez en mesure de prendre des mesures supplémentaires en fonction des ressources, de l'expertise et du temps disponibles. Le document du projet POWRR intitulé *Resources for Technical Steps* (3 pages) ajoute des ressources supplémentaires pour certaines étapes (7 pages).

Kryoflux : Outil commercial pour la lecture de disquettes
http://www.kryoflux.com/
KryoFlux est un appareil basé sur la technologie USB. Il est conçu spécifiquement pour lire de façon fiable à bas niveau et adapté à la préservation des logiciels. Il s'agit d'un matériel officiel développé par la *Software Preservation Society*.

Chamber of Horrors: Obsolete and Endangered Media
http://dpworkshop.org/dpm-eng/oldmedia/disks.html
Quelques exemples de disques obsolètes et menacés.

Lost Formats
http://www.experimentaljetset.nl/archive/lostformats
Page web de la *Lost Formats Preservation Society* donnant un très bel aperçu sous forme de silhouettes des supports, afin de permettre leur identification rapide. Il fournit également un bref historique et les caractéristiques clés de ces supports, comme leurs dimensions et leur capacité de stockage. Toutes les silhouettes sont présentées à la même taille plutôt qu'à l'échelle. La dernière mise à jour majeure semble avoir eu lieu vers 2008, mais le contenu est toujours valable pour tous les supports sauf les plus récents.

Musée des supports obsolètes
http://www.obsoletemedia.org/category/format/
Importante ressource couvrant un très large éventail de supports de stockage obsolètes pour l'audio, la vidéo, les données et les films. Vous pouvez parcourir les catégories ou la Galerie et la Ligne du temps. Particulièrement opportun si vous savez ce que vous cherchez. Tiré en grande partie des entrées pertinentes de Wikipédia.

Études de cas

A Fistful of Floppies: Digital Preservation in Action.
https://ischool.uw.edu/sites/default/files/capstone/posters/JStanley_Capstone_Landscape.pdf
La bibliothèque de l'Université de Washington possède actuellement une petite collection de thèses et de mémoires électroniques qui accompagnent des documents de la fin des années 1980 à 2011 sur disquettes et CD-R. Ces supports atteindront bientôt ou ont déjà dépassé la limite de leur durée de vie prévue. Le projet mené en 2015 a examiné les possibilités de préservation numérique de cette collection de supports en utilisant la criminalistique numérique comme modèle (1 page).

Enford, D., et al 2008, *"Media Matters : developing processes for preserving digital objects on physical carriers at the National Library of Australia",* Papers from 74th IFLA General Conference and Council
http://archive.ifla.org/IV/ifla74/papers/084-Webb-en.pdf
La Bibliothèque nationale d'Australie possédait une collection relativement petite mais importante d'objets numériques sur supports physiques, comprenant à la fois des objets publiés et des manuscrits non publiés sous forme numérique. Le *Digital Preservation Workflow Project* visait à produire un processus semi-automatisé et scalable pour transférer les données des supports physiques vers le stockage de masse numérique de préservation, contribuant à atténuer les principaux risques associés aux supports physiques (17 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
Présentation de la façon de commencer la planification de la préservation numérique, y compris le cadrage, l'évaluation des risques et la manière de prioriser dans votre collection (y compris des exemples de supports anciens), ainsi que les rôles et responsabilités du personnel. 2013 (20 pages).

Références

Brown, A., 2008. Selecting storage media for long-term preservation. *TNA Digital Preservation Guidance Note 2: August 2008*. Disponible à l'adresse suivante : https://www.nationalarchives.gov.uk/documents/selecting-storage-media.pdf
Harvey, R., 2011. *Preserving Digital Materials 2nd edition*. De Gruyter Saur.

Planification de la préservation

Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Qu'est-ce que la planification de la préservation ?

La planification de la préservation est la fonction, au sein d'un entrepôt numérique, qui consiste à surveiller les changements susceptibles d'avoir un impact sur la durabilité d'un objet numérique détenu par l'entrepôt ou sur l'accès à celui-ci. Elle doit être proactive et s'intéresser aux acquisitions et tendances courantes et à venir. Des changements peuvent survenir au sein de l'entrepôt, de l'organisation dont l'entrepôt dépend ou en dehors de l'entrepôt et de l'organisation eux-mêmes. Les changements peuvent être contrôlés dans les domaines suivants :

- Veille technologique
  - Conditionnement
  - Stockage
  - Formats
  - Outils
  - Environnement
  - mécanismes d'accès

- Communautés identifiées
  - besoins et attentes des utilisateurs
  - besoins et attentes des producteurs
  - outils émergents pour un accès de machine à machine
  - retour d'expériences formel de la part des utilisateurs et des producteurs.

Le concept de planification de la préservation est défini dans le modèle fonctionnel de la norme OAIS (CCSDS, 2012). Cette section se concentre principalement sur ce qui permet d'assurer la surveillance au sens de la définition OAIS. Les fonctions « Veille technologique » et « Suivi de la communauté d'utilisateurs cibles » de l'OAIS fournissent des rapports qui informent les activités de planification de la préservation. Elles alertent l'entrepôt sur les changements survenus dans l'environnement externe et les risques qui pourraient avoir un impact sur sa capacité à préserver et à maintenir l'accès aux informations sous sa garde, comme les innovations en matière de technologies de stockage et d'accès, ou les changements dans la portée ou les attentes de la communauté d'utilisateurs cibles (voir Lavoie, 2015,13). La planification de la préservation élabore

ensuite des recommandations pour mettre à jour les politiques et les procédures de l'entrepôt, afin de tenir compte de ces changements. La fonction de planification de la préservation vise à protéger l'entrepôt contre des utilisateurs et des technologies en constante évolution. Elle détecte les changements ou les risques qui ont un impact sur la capacité de l'entrepôt à assumer ses responsabilités, conçoit des stratégies pour y faire face et aide à la mise en œuvre de ces stratégies au sein du système de préservation.

Quel est l'objectif de la planification de la préservation ?

Identifier les événements déclencheurs d'actions de préservation sur les objets numériques

Lorsqu'un changement a été identifié, un processus d'évaluation des risques peut être utilisé pour analyser et identifier le changement qui représente un risque significatif pour l'objet numérique dans l'entrepôt. Les risques peuvent ensuite être traités et, espérons-le, atténués à la suite d'un exercice de planification de la préservation pour décider des mesures de préservation appropriées. Dans ce cas, le processus de surveillance ou de veille technologique permet d'identifier les événements nécessitant de déclencher une analyse plus approfondie, la planification de la préservation et, le cas échéant, les mesures à prendre pour préserver les objets numériques.

Créer une base de connaissances pour alimenter les activités de préservation

Le processus de surveillance des facteurs internes et externes dans le cadre d'une activité de planification de la préservation peut alimenter la base de connaissances d'une organisation et, ce faisant, améliorer sa capacité à mener efficacement des activités de préservation numérique. Par exemple, la « base de connaissances » d'une organisation peut être complétée par des informations sur les capacités d'un nouvel outil logiciel, ou sur l'obsolescence et l'indisponibilité d'un outil existant. Dans certains cas, il vaut mieux mettre en œuvre ce processus à l'échelle individuelle ou organisationnelle, mais, selon les cas, il peut aussi être plus utile de le mettre en œuvre de manière collaborative. La profondeur et l'étendue des connaissances requises pour la préservation numérique favorisent naturellement une approche collaborative, dans laquelle des organisations particulières peuvent se spécialiser dans un domaine particulier et contribuer à une base de connaissances ouverte ou partagée.

Mettre en œuvre un service de planification de la préservation

Le poids de la veille technologique variera en fonction du degré d'uniformité ou de contrôle des formats et des supports qui peut être exercé par l'institution. Celles qui surveillent peu les supports et les formats reçus et conservent des fonds d'une grande diversité trouveront cette fonction essentielle. Dans la plupart des autres institutions, la stratégie SI devrait chercher à développer des normes uniques afin que tout le monde utilise les mêmes logiciels et les mêmes versions, avec des montées vers de nouvelles versions au fur et à mesure du développement des produits.

Si l'on ne met pas en place une veille technologique efficace ou une stratégie SI intégrant cette démarche, on risque de perdre l'accès aux collections numériques et d'augmenter les coûts. Il peut être possible, par exemple, de rétablir l'accès en utilisant les méthodes et outils de la criminalistique numérique (voir Criminalistique numérique), ce qui peut se révéler plus coûteux que des stratégies préventives.

Une étude rétrospective des collections numériques (voir Pour commencer) ainsi qu'une évaluation des risques et un plan d'action (voir Gestion des risques et du changement) peuvent

constituer une première étape nécessaire pour de nombreuses institutions, avant de mettre en place une veille technologique.

De bonnes métadonnées de préservation dans un catalogue informatisé identifiant le support de stockage, le matériel, le système d'exploitation et les logiciels nécessaires permettront de mettre en place une stratégie de veille technologique (voir Métadonnées et documentation).
Les systèmes de préservation intégrés, ainsi que les outils et registres individuels peuvent également prendre en charge cette fonction (voir Solutions et outils techniques).

Ressources

Certaines des principales activités de surveillance de la préservation sont génériques et donc adaptées à une approche collaborative, tandis que d'autres sont très localisées et difficiles à partager.

Série de rapports de veille technologique de la DPC
http://www.dpconline.org/advice/technology-watch-reports
Ces rapports constituent une introduction avancée à des questions spécifiques pour les personnes chargées d'établir ou de gérer des services de préservation et d'accès à long terme. Ils sont mis à jour et de nouveaux rapports sont ajoutés périodiquement.

Scout – a preservation watch system
http://openpreservation.org/blog/2013/12/16/scout-preservation-watch-system/
Ces rapports constituent une introduction avancée à des questions spécifiques pour les personnes chargées d'établir ou de gérer des services de préservation et d'accès à long terme. Ils sont mis à jour et de nouveaux rapports sont ajoutés périodiquement.

Assessing file format risks: searching for Bigfoot?
http://openpreservation.org/blog/2013/09/30/assessing-file-format-risks-searching-bigfoot/
Ce billet de blog détaillé soulève des inquiétudes quant aux défis que pose l'automatisation de la surveillance de la préservation.

Barbara Sierman, Paul Wheatley, “Evaluation of Preservation Planning within OAIS, based on the Planets Functional Model Planets Deliverable no. PP7-D6.1”, 2010.
http://www.planets-project.eu/docs/reports/Planets_PP7-D6_EvaluationOfPPWithinOAIS.pdf[41]
Le projet Planets a concrétisé divers aspects des concepts définis dans le cadre de la fonction de planification de la préservation de l'OAIS, et a procédé à une évaluation de l'OAIS sur la base de ces expériences pratiques. 2010 (34 pages).

[41]NdT : Le document n’était plus accessible en ligne à la date de la présente traduction, une archive est disponible sur : https://web.archive.org/web/20201230142602/http://www.planets-project.eu/docs/reports/Planets_PP7-D6_EvaluationOfPPWithinOAIS.pdf

Community Owned digital Preservation Tool Registry COPTR
http://coptr.digipres.org/Main_Page
Le registre COPTR décrit des outils utiles pour la préservation numérique à long terme et agit principalement comme un outil de recherche et d'évaluation pour aider les professionnels à trouver les outils dont ils ont besoin pour préserver les données numériques. Le registre COPTR vise à rassembler en un seul endroit les connaissances de la communauté de la préservation numérique sur les outils de préservation. Il a été initialement alimenté par des données provenant de registres gérés par les organisations partenaires du registre COPTR, notamment ceux tenus par le *Digital Curation Centre*, le *Digital Curation Exchange*, la *National Digital Stewardship Alliance*, l'*Open Preservation Foundation* et le *Preserving digital Objects With Restricted Resources project* (POWRR). Le registre COPTR permet de saisir des détails factuels de base sur un outil, ses fonctions, la manière de trouver plus d'informations (URL pertinentes) et des références aux expériences des utilisateurs avec l'outil. Le champ d'application est une interprétation large du terme « préservation numérique ». En d'autres termes, si un outil est utile pour effectuer une fonction de préservation numérique telle que celles décrites dans le modèle OAIS ou le modèle de cycle de vie DCC, alors il entre dans le champ d'application de ce registre.

Études de cas

Rapport de recherche d'OCLC – Preservation Health Check : Surveillance des menaces pesant sur le contenu des dépôts numériques
http://www.oclc.org/research/themes/research-collections/phc.html
L'activité *Research Preservation Health Check* d'OCLC a été initiée par l'*Open Planets Foundation*. Le pilote a utilisé un échantillon de métadonnées de préservation fournies par la Bibliothèque nationale de France. Le rapport présente les résultats préliminaires de la phase 1 du pilote et suggère qu'il est possible d'utiliser les métadonnées de préservation PREMIS comme base de preuves pour réaliser un exercice d'évaluation des menaces basé sur le modèle SPOT (*Simple Property- Oriented Threat*).

Digital Preservation planning. Case Study.
http://www.dpconline.org/component/docman/doc_download/863-2013-may-getting-started-london-planning-case-study-ed-fay
Présentation sur la manière de commencer la planification de la préservation numérique, y compris le cadrage, l'évaluation des risques et la manière de prioriser dans votre collection (y compris des exemples de médias anciens), ainsi que les rôles et responsabilités du personnel (2013, 20 pages).

Références

Consultative Committee for Space Data Systems, 2012. *Reference model for an open archival information system (OAIS): Recommended practice* (CCSDS 650.0-M-2: Magenta Book), CCSDS, Washington, DC. Disponible à l'adresse suivante : https://public.ccsds.org/pubs/650x0m2.pdf
(Il s'agit d'une version librement téléchargeable de l'ISO 14721:2012 – Systèmes de transfert des informations et données spatiales – Système ouvert d'archivage d'information (OAIS) – Modèle de référence, 2e édition).

Version française disponible à l'adresse suivante :
https://public.ccsds.org/Pubs/650x0m2%28F%29.pdf

Lavoie, B., 2014. *The Open Archival Information System (OAIS) Reference Model: Introductory Guide (2nd Edition)* DPC Technology Watch Report 14-02 October 2014. Disponible à l'adresse suivante : http://dx.doi.org/10.7207/TWR14-02

Mesures de préservation

Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Introduction

Nous savons maintenant que la préservation numérique consiste à relever une série de défis liés à des questions d'organisation, de ressources, de gestion, de culture et de technique. Cette section du manuel se concentrera spécifiquement sur les mesures qui peuvent être prises pour aider à atténuer les défis techniques de la préservation des objets numériques au fil du temps.

Obsolescence technologique des formats

L'obsolescence technologique a longtemps été considérée comme un défi important de la préservation numérique à long terme. Cependant, ces dernières années, des études ont suggéré que l'obsolescence des formats n'est pas toujours aussi fréquente qu'on le craignait (Rosenthal 2015a, Jackson 2012). C'est un problème qui doit être identifié et combattu si l'on veut que les objets numériques survivent à des générations de changements technologiques, mais ce n'est certainement pas le seul problème. De nombreux formats de fichiers établis existent toujours, sont toujours pris en charge et toujours utilisables. Il est fort probable que la majorité des formats de fichiers que vous traitez seront communément compris et bien pris en charge plutôt qu'obsolètes.

On peut définir simplement l'obsolescence comme le processus qui consiste à devenir dépassé ou à ne plus être utilisé. Lorsque nous parlons d'obsolescence technologique, nous faisons par exemple référence à « ce logiciel Wordperfect 3.1 est obsolète » ou « cet ordinateur BBC Micro est obsolète ». Le moment exact auquel l'obsolescence se produit peut-être difficile à déterminer, en particulier pour les matériels qui sont devenus obsolètes récemment. Par exemple, ce n'est pas parce que l'application d'origine (par exemple MS Word) ne prend plus en charge un format donné qu'aucun autre logiciel capable de lire ce format n'est pas disponible. De même, une institution peut continuer à utiliser et à maintenir un logiciel ancien longtemps après que d'autres aient mis à jour leurs logiciels vers de nouvelles versions. Il est donc peut-être plus exact de parler d'« obsolescence institutionnelle », c'est-à-dire que la technologie en question n'est plus utilisée ou facilement accessible par une institution donnée.

L'obsolescence est un problème car tous les fichiers ont leurs propres dépendances matérielles et logicielles. C'était particulièrement le cas dans les premiers temps de l'informatique.

Le changement devient un problème lorsqu'il compromet la signification du contenu ou son interprétation par un utilisateur. Un objectif essentiel des mesures de préservation numérique est de préserver l'intégrité et l'authenticité de l'objet conservé, malgré ces changements générationnels dans la technologie informatique. Dans la section suivante, nous examinerons quelques stratégies communes pour aider à minimiser ces changements.

Stratégies de préservation

Dans cette section, nous passons en revue les stratégies techniques qui peuvent être employées pour préserver l'information numérique. Après une première vague à la fin des années 1990, la recherche de nouvelles stratégies a relativement peu progressé, bien que des travaux de recherche et de développement importants aient été menés sur les différentes options de mise en œuvre et les technologies de soutien, telles que l'assurance qualité, la criminalistique numérique (voir la section criminalistique numérique) et les registres d'informations de représentation technique (voir la section Solutions et outils techniques du manuel). Les techniques que nous allons aborder ici sont les suivantes :

- Migration des formats
- Émulation
- Musées de l'informatique

Migration des formats

La migration de format est l'une des stratégies de préservation les plus utilisées et la plupart des systèmes de préservation numérique contiennent des fonctionnalités ou des données système qui supposent une solution de migration. La migration des formats est différente de la migration des supports de stockage. Elle implique le transfert ou la transformation (c'est-à-dire la migration) de données d'un format vieillissant ou obsolète vers un nouveau format, en utilisant éventuellement de nouveaux systèmes d'application à chaque étape pour interpréter les informations. Le passage d'une version d'une norme de format à une norme ultérieure est un aspect de cette méthode ; par exemple, le passage de la version 6 de MS Word (de 1993) à MS Word pour Windows 2010. Pour des cadres et des outils utiles pour évaluer l'obsolescence technique des formats de fichiers, voir Formats de fichiers et normes.

La migration de format, comme toute intervention susceptible de modifier la structure et le contenu des données, peut introduire des erreurs et des pertes d'informations. Il est donc important de définir des indicateurs permettant de mesurer la perte éventuelle d'informations et de les utiliser pour évaluer au moyen de tests l'exactitude et la qualité de la migration de format.
Les travaux récents concernant l'assurance qualité et les mesures de préservation numérique comprennent les travaux des projets AQUA, SPRUCE et SCAPE. Pour mesurer les taux d'erreur, il est nécessaire de déterminer certains indicateurs très spécifiques. Vous devrez peut-être définir ce que vous considérez comme une erreur et si vous considérez que certaines erreurs sont plus importantes que d'autres. Cela dépend du contexte/contenu du document et des caractéristiques des objets jugés importants à préserver, ainsi que des outils de migration et des formats successifs utilisés dans tout processus de migration.

Parmi les questions pratiques qui se posent dans ce processus, on peut citer le choix du moment de la migration : vaut-il mieux migrer de génération en génération, ou faut-il sauter certaines générations ? Vous devrez enregistrer toutes les transformations, leurs résultats et documenter les pertes d'informations détectées afin de conserver des preuves d'authenticité et d'autorité. PREMIS peut être un outil utile à cet effet - voir la section du manuel sur Métadonnées et documentation pour plus d'informations sur ce standard. Il est de bonne pratique de toujours conserver le format de fichier original tel qu'il a été pris en charge pour y revenir si nécessaire.

Émulation

L'émulation offre une solution alternative à la migration qui permet aux entrepôts numériques de pérenniser et de donner accès aux utilisateurs directement aux fichiers originaux. Cette technique tente de préserver les comportements originaux et l'aspect des applications, ainsi que le contenu informationnel. Elle repose sur l'idée que seul le programme original fait autorité en matière de format, ce qui est particulièrement utile pour les objets complexes présentant de multiples interdépendances, tels que les jeux ou les applications interactives.

Un émulateur, comme son nom l'indique, est un programme qui fonctionne sur une architecture informatique courante mais qui offre les mêmes possibilités et le même comportement qu'une architecture ancienne. Cette approche a été approuvée par un certain nombre d'organisations patrimoniales, souvent en collaboration avec des experts techniques et, ces dernières années, des succès notables ont été enregistrés dans la mise en œuvre de solutions d'émulation pour le patrimoine culturel (voir Ressources ci-dessous). Toutefois, il reste des défis importants à relever, notamment les problèmes de droits associés aux licences de logiciels qui doivent souvent être résolus (Rosenthal 2015b).

Un avantage particulier de l'émulation est qu'une seule solution peut être déployée pour donner accès à un grand nombre d'objets, à condition que tous ces objets soient livrés sur le même système d'exploitation ou le même empilement matériel. L'utilisation d'un équipement informatique ancien peut toutefois s'avérer difficile pour les utilisateurs, bien qu'ils aient presque certainement accès à une représentation authentique des documents. Bien entendu, les émulateurs doivent être construits et entretenus, ce qui nécessite la mise à disposition d'un groupe d'experts, et ne peut pas toujours être faisable. De nouveaux émulateurs seront nécessaires à mesure que les architectures informatiques deviendront obsolètes, et ces deux éléments entraînent des coûts et des besoins en ressources.

Musées de l'informatique

Cette méthodologie propose la préservation des ordinateurs et de leurs logiciels système (systèmes d'exploitation, pilotes, etc.) ainsi que des données et des programmes d'application. Des efforts doivent être déployés pour maintenir toutes les plateformes en bon état et pour conserver toutes les connaissances nécessaires à la maintenance et à l'utilisation des machines et de leurs programmes. L'idée repose également sur l'existence d'un stock de pièces de rechange, mais celles-ci vont se réduire, tout comme les groupes d'experts. C'est pourquoi cette stratégie tend à être une mesure provisoire plutôt qu'une solution à long terme. Il existe quelques musées officiels, comme le *Computer History Museum* en Californie et le *Centre for Computing History* à Cambridge. Ils maintiennent généralement les machines en état de marche, mais n'offrent pas de services de préservation. Voir également la section Anciens supports du manuel pour de plus amples informations sur les formats de fichiers et les supports historiques.

Mise en œuvre

Les rapports de veille technologique de la DPC constituent un guide particulièrement utile sur les typologies et les formats de fichiers les plus courants (y compris les courriers électroniques, les réseaux sociaux, l'audiovisuel, les livres électroniques, les périodiques électroniques, les SIG, la CAO, l'archivage Web, etc.) et montrent quelles stratégies tendent à être utilisées pour chaque catégorie. Les outils d'aide à la mise en œuvre des stratégies de préservation sont abordés dans la partie du manuel consacrée aux solutions et outils techniques, notamment en ce qui concerne les formats de fichiers et les normes.

Ressources

Série de rapports de veille technologique de DPC

http://www.dpconline.org/publications/technology-watch-reports

La série de rapports de veille technologique de la DPC est conçue comme une introduction avancée à des questions spécifiques pour ceux qui sont chargés d'établir ou de gérer des services d'accès à long terme. Ils identifient et suivent l'évolution des technologies informatiques, des normes et des outils qui sont essentiels aux activités de préservation numérique. Ils sont commandés par des experts sur ces questions et sont minutieusement évalués par des pairs avant d'être publiés.

Emulation and Virtualization as Preservation Strategies

https://mellon.org/media/filer_public/0c/3e/0c3eee7d-4166-4ba6-a767-6b42e6a1c2a7/rosenthal-emulation-2015.pdf

Ce rapport sur l'émulation et la virtualisation en tant que stratégies de préservation, publié en 2015 et rédigé par David Rosenthal, a été financé par la Fondation Mellon, la Fondation Sloan et l'IMLS. Il conclut que les récents développements des cadres d'émulation permettent de fournir des émulations aux lecteurs via le Web de manière à ce qu'ils apparaissent comme des composants normaux des pages Web. Cela supprime ce qui était le principal obstacle au déploiement de l'émulation en tant que stratégie de préservation. Les deux obstacles les plus importants qui subsistent sont les suivantes : les outils de création d'images système à préserver sont inadéquats ; le cadre juridique pour la fourniture d'émulations n'est pas clair, et lorsqu'il est clair, il est très restrictif. Ces deux obstacles augmentent le coût de la construction et de l'accès à une collection substantielle et bien conservée d'objets numériques émulés, ce qui le rend hors de portée. Si ces obstacles peuvent être levés, l'émulation jouera un rôle beaucoup plus important dans la préservation numérique dans les années à venir (37 pages).

Becker, C. et al., “Systematic planning for digital preservation: evaluating potential strategies and building preservation plan”, 2009.

http://www.ifs.tuwien.ac.at/~becker/pubs/becker-ijdl2009.pdf

Cet article publié en 2009 décrit une approche systématique pour évaluer les alternatives potentielles aux mesures de préservation et pour élaborer des plans de préservation bien définis et responsables pour maintenir le contenu numérique en vie au fil du temps. Le travail a été entrepris dans le cadre du projet PLANETS financé par l'Union européenne (25 pages).

File Format Conversion
http://www.nationalarchives.gov.uk/documents/information-management/format-conversion.pdf
La conversion de format peut vous aider à maintenir l'accès et l'utilisation de vos informations et à atténuer les risques liés à l'obsolescence. Cette note d'information des Archives nationales du Royaume-Uni de 2011 vous indique les étapes à suivre pour mettre en œuvre une opération de conversion de format de fichier(29 pages).

What organizations are preserving software?
http://qanda.digipres.org/1068/what-organizations-are-preserving-software
Ce billet et les réponses fournies à partir d'août 2015 sur le site *Digital Preservation Q&A* fournit une liste et des liens utiles pour les institutions préservant des logiciels afin de mettre en œuvre des stratégies d'émulation.

Projet SCAPE Lignes directrices et recommandations finales sur les meilleures pratiques
http://scape-project.eu/wp-content/uploads/2014/02/SCAPE_D20.6_KB_V1.0.pdf[42]
Ce rapport du projet SCAPE publié en 2014 couvre trois domaines principaux : la mise en œuvre de la migration à grande échelle comme stratégie de préservation. Les autres domaines sont la préservation des données de recherche et la préservation du train de bits. (127 pages).

Études de cas

Internet Arcade
https://archive.org/details/internetarcade
L'Internet Arcade est une bibliothèque en ligne de jeux vidéo d'arcade (à pièces) des années 1970 aux années 1990 de l'Internet Archive, utilisant une solution d'émulation intégrée au navigateur pour donner accès à la collection.

Rhizome
https://sites.rhizome.org/theresa-duncan-cdroms/
Dans les années 1990, Theresa Duncan et ses collaborateurs ont réalisé trois jeux vidéo qui illustrent le meilleur de la narration interactive. Deux décennies plus tard, ces œuvres (comme la plupart des CD-ROM) sont tombées dans l'oubli. Cette exposition en ligne, co-présentée par Rhizome et le New Museum leur redonne une nouvelle vie, les rendant jouables en ligne par émulation.

Fritz, C. et al., “View of Assessing Migration Risk for Scientific Data Formats”, 2012
http://www.ijdc.net/index.php/ijdc/article/view/202/271
Ce document explore une hypothèse simple - selon laquelle, lorsque des voies de migration existent, la majorité des fichiers de données scientifiques peuvent être migrées en toute sécurité, n'en laissant que quelques-uns qui doivent être traités avec plus de précaution - dans un contexte où plusieurs formats de données scientifiques sont ou étaient largement utilisés. L'approche

[42] NdT : Le document n’était plus accessible en ligne à la date de la présente traduction, une archive est disponible sur : https://web.archive.org/web/20201229173233/https://www.scape-project.eu/wp-content/uploads/2014/02/SCAPE_D20.6_KB_V1.0.pdf

consiste à recueillir des informations sur les éventuelles inadéquations de migration et, à l'aide d'outils personnalisés, à évaluer l'incidence de ces risques sur une vaste collection de fichiers de données. Les résultats confirment l'hypothèse initiale, mais avec quelques réserves.

Portico – *Preservation step by step*
https://www.portico.org/our-work/preservation-step-step/
Un guide utile, étape par étape, sur la planification de la préservation et les stratégies de migration employées par Portico. Le plan de préservation peut inclure une migration initiale du conditionnement ou des fichiers dans des formats spécifiques (par exemple, Portico migre les articles de périodique électronique au format XML vers la norme d'archivage NLM).

Thorpe, C.,"*Trash to Teasure: Retro computer, software collection helps National Library access digital pieces*", ABC News, 2015
http://www.abc.net.au/news/2015-06-20/collecting-retro-computer-technology-to-save-digital-treasures/6560494
La Bibliothèque nationale d'Australie a rendu publics ses propres efforts pour développer une collection de matériels et de logiciels informatiques anciens. Elle l'utilise pour faciliter la récupération des données et met ensuite en œuvre d'autres stratégies de préservation et ne compte pas sur le musée de l'informatique pour la préservation à long terme.

Références

David Rosenthal, 2015a. *"The Prostate Cancer of Preservation" Re-examined.* Disponible à l'adresse suivante : http://blog.dshr.org/2015/09/the-prostate-cancer-of-preservation-re.html

David Rosenthal, 2015b. *Emulation & Virtualization as Preservation Strategies*. Disponible à l'adresse suivante : https://mellon.org/media/filer_public/0c/3e/0c3eee7d-4166-4ba6-a767-6b42e6a1c2a7/rosenthal-emulation-2015.pdf

Andrew N. Jackson, 2012. *Formats over Time: Exploring UK Web History*. Disponible à l'adresse suivante : http://arxiv.org/abs/1210.1714

Accès

Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Introduction

Il y a toujours eu un lien étroit entre la préservation et l'accès. L'objectif principal de la préservation du contenu informationnel des ressources traditionnelles est de faire en sorte qu'il reste accessible aux générations actuelles et futures. Préserver l'accès aux objets numériques est l'objectif principal des programmes de préservation numérique, mais cela nécessite une gestion active tout au long du cycle de vie de la ressource avant de pouvoir être assuré. Il est donc essentiel de prendre en considération les problématiques de fourniture de l'accès dès le début du processus de préservation, idéalement dès la phase d'acquisition. Ceci est représenté dans l'Arbre de décision pour la sélection des objets numériques pour la préservation à long terme, inclus dans la section Acquisition et évaluation de ce manuel. Dans cette optique, cette section vise à identifier les principales questions qui doivent être prises en compte, les décisions qui doivent être prises lors de la planification de la fourniture d'accès et la manière dont celle-ci peut avoir un impact sur la préservation de manière plus générale.

Comprendre les utilisateurs

La compréhension des utilisateurs potentiels est essentielle pour la planification de l'accès aux objets numériques et constitue un élément clé des activités de préservation numérique au sens large. L'importance de ce travail est peut-être plus évidente dans l'accent mis sur les « communautés d'utilisateurs cible » dans le Modèle de référence pour un système d'archivage ouvert (OAIS). Les connaissances acquises sur ces utilisateurs cibles éclaireront les décisions prises tout au long du cycle de vie, mais elles auront probablement le plus de poids lors du choix des solutions de fourniture d'accès appropriées, en tenant compte des ressources et des considérations technologiques. Il est important d'aborder l'identification des communautés d'utilisateurs cible et de leurs besoins de manière systématique et objective. En bref, il faut comprendre ce que les utilisateurs veulent faire et quelles fonctionnalités peuvent être fournies par l'entrepôt numérique.

La méthodologie utilisée pour la collecte de ces informations variera en fonction du contexte organisationnel. Les options et outils potentiels peuvent comprendre les éléments suivants :

- Analyse de l'utilisation actuelle (demandes d'accès aux objets physiques et numériques, statistiques du site web, etc.)

- Enquêtes
- Groupes de discussion
- Interviews
- Recueil de cas d'usage
- Analyse des tâches

Lors de l'analyse de l'usage, il est important de prendre en compte à la fois les utilisateurs existants et les non-utilisateurs. Bien que l'interaction avec les non-utilisateurs soit par nature plus difficile, elle peut être un processus utile pour comprendre les obstacles actuels à l'utilisation ainsi que pour identifier de nouveaux segments de marché potentiels.

Une fois collectées, ces informations doivent être utilisées pour éclairer les décisions prises en ce qui concerne la mise en œuvre des solutions de fourniture d'accès. Il est également important de continuer à surveiller le développement des communautés d'utilisateurs, qu'il convient d'intégrer aux activités de planification de la préservation au sein de votre organisation.

Formats d'accès

Le format dans lequel les objets numériques seront livrés aux utilisateurs est un élément clé à prendre en compte lors de la planification de l'accès. S'il existe un lien étroit entre la préservation et l'accès en ce qui concerne l'objectif primordial d'un programme de préservation numérique, il est également nécessaire d'établir une distinction claire entre les deux. Il peut y avoir une combinaison de raisons techniques, juridiques et pragmatiques pour séparer la version de diffusion de la version de préservation, de sorte qu'il peut être souhaitable, voire nécessaire, de livrer une version de diffusion de l'objet numérique à l'utilisateur dans un format différent de celui qui est conservé dans le système de préservation. En effet, une organisation peut souhaiter offrir différentes « saveurs » de format en fonction des besoins de l'utilisateur ou de la communauté en question. Lors de la sélection des formats d'accès, une organisation devra tenir compte de plusieurs questions, dont les suivantes :

- Quel est le format le plus couramment utilisé/largement pris en charge pour le type d'objet ?
- Les utilisateurs auront-ils accès à des visionneuses/logiciels gratuits qui prennent en charge le type de fichier proposé ?
- Quelle est la taille du fichier produit et quelles sont les implications pour sa livraison à l'utilisateur ?
- Le format est-il facile à utiliser ?
- Les utilisateurs auront-ils besoin de conseils ou de documents d'appui pour leur permettre d'accéder aux objets ou de les utiliser ?
- L'organisation dispose-t-elle de communautés d'utilisateurs distinctes ayant des besoins différents en matière d'accès ?

Voir aussi Formats de fichiers et normes pour plus de détails sur les formats de préservation et d'accès communs.

Questions juridiques relatives à l'accès

Toute une série de questions juridiques différentes devront probablement être abordées lors de la fourniture de l'accès aux objets numériques. Elles affecteront à la fois les solutions technologiques déployées ainsi que les personnes qui peuvent accéder à l'objet et le moment où elles le font. Les questions juridiques sont l'une des principales considérations en matière d'accès qui se recoupe avec les acquisitions, comme mentionné ci-dessus, et il est essentiel que les informations correctes soient recueillies à ce moment pour faciliter l'accès plus tard dans le cycle de vie. Sans ces

informations, il peut être impossible de gérer correctement l'accès et l'organisation peut s'exposer à un certain nombre de risques juridiques potentiels.

Les questions juridiques à prendre en compte sont notamment les suivantes :

- Restrictions d'utilisation liées à la sensibilité et à la protection des données
- Délais de rétention acceptés sur le contenu lorsque l'accès anticipé peut constituer une rupture de contrat
- Gestion des droits de propriété intellectuelle, par exemple le droit d'auteur.

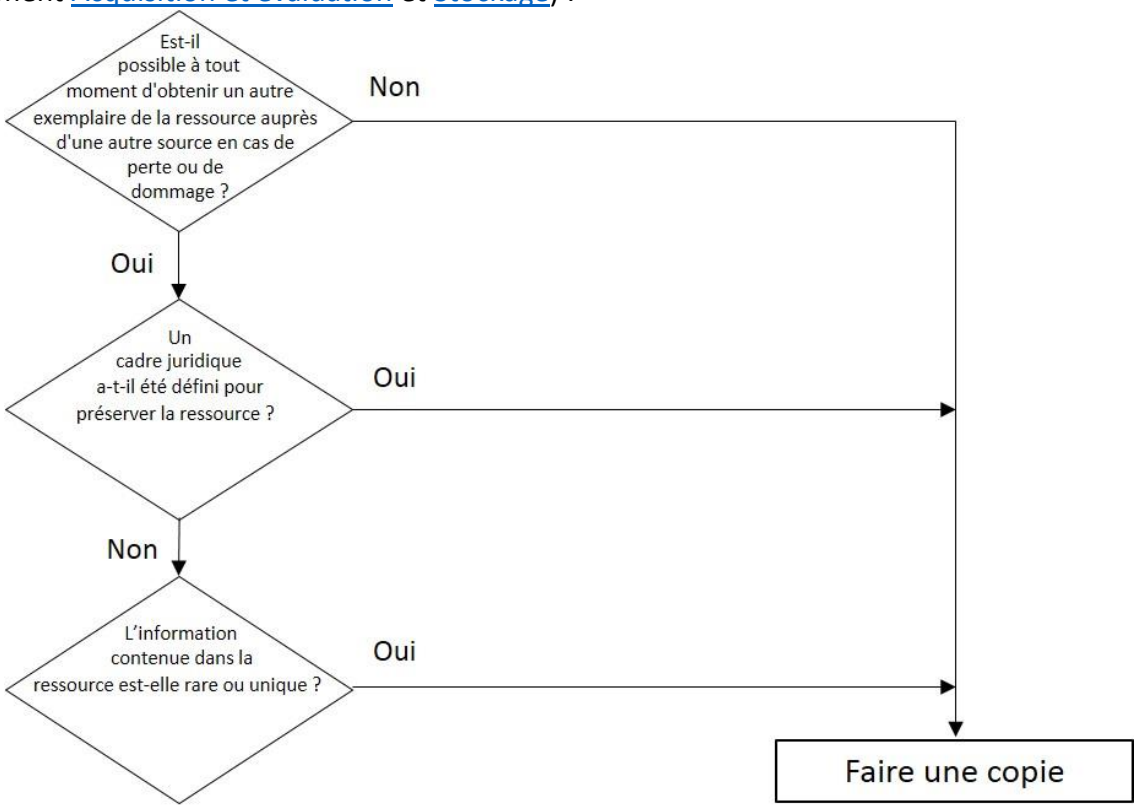
La gestion des droits de propriété intellectuelle en particulier devrait être intégrée dans le processus d'acquisition en prenant soigneusement en considération les accords de transfert et de propriété et les licences de droits d'auteur mis en place à ce moment-là. Les licences doivent clairement indiquer les autorisations d'accès et de réutilisation, y compris les licences de tiers. Celles-ci doivent ensuite être clairement décrites dans la politique et les procédures d'accès, qu'elles soient gérées dans un système de gestion des droits ou selon d'autres méthodes.

Formes de fourniture d'accès

Les décisions clés finales qu'une organisation doit prendre se présentent sous la forme de

- Politique
- Procédure
- Services gratuits ou payants
- Accès en ligne/hors ligne, et environnement d'accès fourni
- Accès pour les personnes en situation de handicap
- Stockage et sécurité

Si la version de diffusion est la seule version d'une ressource numérique, alors le risque de perte par vol ou dommage est clairement très élevé. Si cette approche est adoptée, il convient de procéder à une évaluation des risques comprenant certaines des questions suivantes (voir également Acquisition et évaluation et Stockage) :



Conclusions

L'accès est étroitement lié à de nombreuses autres questions et technologies de préservation numérique couvertes dans le manuel. En particulier, vous aurez peut-être intérêt à vous pencher sur les sections Politiques et stratégies institutionnelles, Conformité juridique, Métadonnées et documentation, Acquisition et évaluation, Stockage, Anciens supports, Formats de fichiers et normes et la Sécurité de l'information.

Ressources

Appel, R. et al. "Born-Digital Access in Archival Repositories: Mapping the Current Landscape", 2015.
https://docs.google.com/document/d/15v3Z6fFNydrXcGfGWXA4xzyWlivirfUXhHoqgVDBtUg/preview?sle=true#
Ce document intéressant présente les résultats préliminaires et l'analyse d'une étude et d'une enquête sur les pratiques actuelles d'accès aux œuvres numériques dans plus de 200 institutions du patrimoine culturel. Les répondants étaient principalement originaires des États-Unis.
csds.org/pubs/650x0m2.pdf

OAIS : *Open Archival Information Systems* : Modèle de référence pour un système d'information archivistique ouvert: Pratique recommandée (CCSDS 650.0-M-2 : livre magenta), Comité consultatif pour les systèmes de données spatiales 2012
https://public.ccsds.org/pubs/650x0m2.pdf (version anglaise)
Fournit une terminologie et un modèle fonctionnel partagés utiles pour identifier les exigences relatives à l'acquisition de services de préservation numérique par des tiers.
Il a ensuite été publié sous la forme de la norme ISO 14721:2012 – Systèmes de transfert des informations et données spatiales – Système ouvert d'archivage d'information (OAIS) – Modèle de référence, 2e édition. La fonction d'accès au sein de l'OAIS gère les processus et les services par lesquels les consommateurs - et en particulier la communauté d'utilisateurs cible - localisent, demandent et reçoivent livraison des articles gérés par la fonction stockage de l'OAIS. En tant que tel, il s'agit du principal mécanisme par lequel les entrepôts numériques s'acquittent de leur responsabilité de mettre leurs informations à la disposition de leur communauté d'utilisateurs. (135 pages).

Version française : https://public.ccsds.org/Pubs/650x0m2%28F%29.pdf

Brown, A., *Practical Digital Preservation a how-to guide for organizations of any size*, 2013.
Le chapitre 9 (28 pages) de ce livre est consacré à la question de l'accès des utilisateurs.

Registre des outils de préservation numérique appartenant à la communauté COPTR

http://www.digipres.org/tools/

Il existe un grand nombre d'outils d'accès ou qui comportent des fonctionnalités d'accès. Le manuel recommande de les rechercher via l'outil POWRR Grid dans COPTR. La grille d'outils POWRR fournit un ensemble de vues interactives conçues pour aider les professionnels à identifier et à sélectionner les outils dont ils ont besoin pour résoudre les problèmes de préservation numérique. La colonne Accès, Utilisation et Réutilisation de la grille identifie les outils d'accès pour des types de contenu spécifiques ou des outils et systèmes génériques qui ont des fonctions d'accès. Des hyperliens sont partout présents dans la grille, il suffit donc de cliquer sur ceux-ci jusqu'à ce que vous trouviez l'information que vous recherchez. En cliquant sur le nom d'un outil de préservation spécifique, vous obtiendrez plus de détails sur le wiki de COPTR, où vous devez vous rendre pour développer ou mettre à jour les informations sur les outils.

AIMS Born-Digital Collections: An Inter-Institutional Model for Stewardship

http://dcs.library.virginia.edu/files/2013/02/AIMS_final.pdf

Le cadre AIMS (*An Inter-Institutional Model for Stewardship*) est une méthodologie pour la gestion des objets numériques natifs. Il est divisé en quatre sections principales, correspondant aux meilleures pratiques de haut niveau pour les processus de travail des documents numériques : développement de la collection, collecte, classement et description, recherche et accès. L'accès se concentre principalement sur la rédaction et les informations sensibles. Les annexes comprennent, par exemple, des diagrammes de processus de travail pour le traitement des échantillons, une analyse des outils et des enquêtes auprès des donateurs (195 pages).

Études de cas

Études de cas des Archives nationales du Royaume-Uni : Accès en ligne
http://www.nationalarchives.gov.uk/archives-sector/online-access.htm
Une série de neuf études de cas publiées par les Archives nationales du Royaume-Uni sur la façon dont les collections ont été rendues plus accessibles par la mise en ligne des documents. Elles sont tirées d'une grande variété d'entrepôts numériques.

Casseurs de codes : les créateurs de la génétique moderne
https://digirati.com/work/galleries-libraries-archives-museums/case-studies/wellcome-library/
Une étude de cas réalisée par digirati, les développeurs du visualiseur de la *Wellcome Trust Library*, qui se concentre sur l'utilisation de ce visualiseur pour accéder à la collection Francis Crick. Le visualiseur peut être utilisé pour afficher tous les types de contenu numérique, y compris les livres d'un bout à l'autre, les archives, les œuvres d'art, les vidéos et les fichiers audio. Le logiciel peut être téléchargé sur le compte GitHub de la *Wellcome Library* (https://github.com/wellcomelibrary/player).

Light, M., "Managing Risk with a Virtual Reading Room: Two Born Digital Projects", 2014
http://digitalscholarship.unlv.edu/cgi/viewcontent.cgi?article=1462&context=lib_articles
Entre 2010 et 2013, l'Université de Californie, Irvine, a lancé un site permettant d'accéder en ligne aux documents personnels de Richard Rorty et Mark Poster sous la forme d'une salle de lecture virtuelle. La salle de lecture virtuelle a permis d'atténuer les risques liés à ce type d'accès à des documents personnels présentant des problèmes de confidentialité et de droits d'auteur, en limitant le nombre d'utilisateurs qualifiés et en limitant les fonctionnalités de recherche en plein texte sur le web. L'étude de cas passe en revue chaque phase de recherche et de réflexion, y compris les projets comparables menés dans d'autres institutions et les enseignements tirés de manière très ouverte et informative.

Schein, C., "From Accession to Access: A Born-Digital Materials Case Study", 2014
http://digitalcommons.usu.edu/cgi/viewcontent.cgi?article=1036&context=westernarchives
Entre 2011 et 2013, le *Getty Institutional Records and Archives* a fait sa première incursion dans le domaine de l'entrée, du classement, de la description et de la livraison des objets numériques uniques, lorsqu'il a reçu des archives orales générées par certains des partenaires du projet *Pacific Standard Time : Art in L.A*. Cette étude de cas aborde les défis et les moyens inhérents à cette collection hybride d'enregistrements audiovisuels, de fichiers numériques à supports mixtes et de transcriptions analogiques. Elle décrit les efforts du service d'archives pour développer un processus de traitement générique qui applique la stratégie de gestion des ressources communément appelée MPLP dans un environnement numérique, tout en s'efforçant de sauvegarder l'intégrité et l'authenticité des fichiers, d'adhérer aux normes professionnelles et de respecter les principes archivistiques fondamentaux. L'étude décrit le processus de travail qui en est résulté et met en évidence quelques-unes des technologies peu coûteuses qui ont été utilisées avec succès pour automatiser ou accélérer les étapes du traitement des contenus qui ont été transférés via des supports facilement accessibles et qui se composaient de formats de fichiers courants.

Métadonnées et documentation

Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Introduction

Cette section donne un bref aperçu sur les métadonnées et la documentation de niveau débutant à intermédiaire, en mettant l'accent sur le standard PREMIS de métadonnées de préservation numérique. Elle s'appuie sur la 2e édition du rapport de la veille technologique de la DPC sur les métadonnées de préservation. Le rapport lui-même aborde un plus grand nombre de questions et de pratiques de manière plus approfondie, avec des lectures et des conseils supplémentaires (Gartner et Lavoie, 2013). Il est recommandé aux lecteurs qui ont besoin d'informations à un niveau plus avancé.

Les métadonnées sont des données relatives aux ressources numériques, stockées sous une forme structurée adaptée au traitement automatisé. Elles servent à de nombreux usages dans la préservation à long terme, en fournissant un enregistrement des activités qui ont été réalisées sur les objets numériques et une base sur laquelle les décisions futures sur les activités de préservation peuvent être prises à l'avenir, ainsi qu'en facilitant la recherche et l'accès. Les informations contenues dans un enregistrement de métadonnées englobent souvent toute une série de sujets. Il n'y a pas de ligne de démarcation claire entre ce qui constitue des métadonnées de préservation et ce qui n'en constitue pas, mais en fin de compte, l'objectif des métadonnées de préservation est de rendre possibles les objectifs de la préservation numérique à long terme, qui sont de maintenir la disponibilité, l'identité, la persistance, la capacité de rendu, l'intelligibilité et l'authenticité des objets numériques sur de longues périodes.

La documentation est l'information (telle que les manuels de logiciels, la conception des enquêtes et les guides d'utilisation) fournie par un créateur et l'entrepôt numérique, qui complète les métadonnées et fournit suffisamment d'informations pour permettre l'utilisation de la ressource par d'autres. C'est souvent le seul objet qui donne un aperçu de la manière dont une ressource numérique a été créée, manipulée, gérée et utilisée par son créateur et c'est souvent la clé qui permet aux autres d'utiliser la ressource en connaissance de cause.

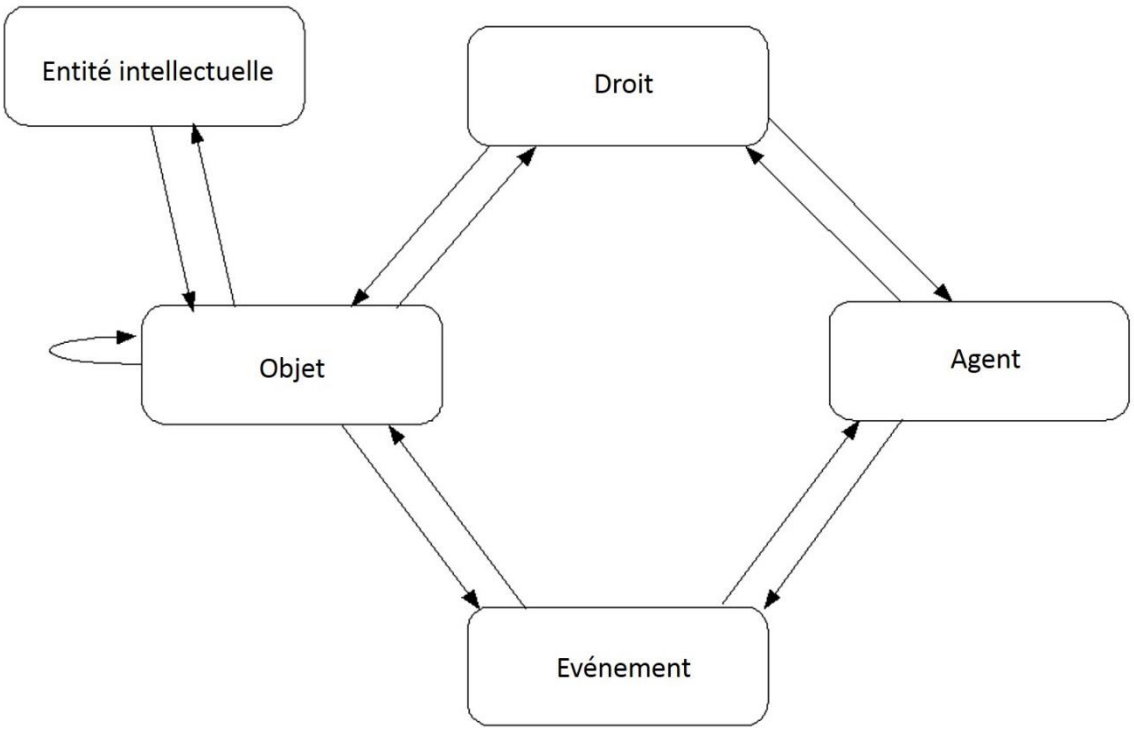
Un certain nombre de facteurs rendent les métadonnées et la documentation particulièrement critiques pour garantir la viabilité des objets numériques et ils sont liés aux différences fondamentales entre les ressources traditionnelles et numériques :

- **Technologie**. Les ressources numériques dépendent du matériel et des logiciels pour les rendre intelligibles. Les exigences techniques doivent être enregistrées afin de pouvoir prendre des décisions sur les stratégies de préservation et d'accès appropriées.
- **Changement**. Alors que les objets traditionnels peuvent être préservés par des programmes de préservation préventive essentiellement passifs, les objets numériques feront l'objet d'actions répétées, et il y aura de nombreux intervenants différents et très probablement différentes institutions qui influenceront la gestion des objets numériques sur une période prolongée. L'enregistrement des actions entreprises sur une ressource et des changements qui en découlent fournira une clé aux futurs gestionnaires et usagers de la ressource.
- **Authenticité**. Les métadonnées et la documentation peuvent être les principaux, sinon les seuls, moyens d'attester de manière fiable de l'authenticité de l'objet à la suite de changements.
- **Gestion des droits**. Si les ressources traditionnelles peuvent ou non être dupliquées dans le cadre de leur programme de préservation, les ressources numériques doivent l'être si l'on veut qu'elles restent accessibles. Les gestionnaires doivent savoir qu'ils ont le droit de dupliquer à des fins de préservation, quelles technologies (le cas échéant) ont été utilisées pour contrôler la gestion des droits et quelles sont les implications (le cas échéant) pour contrôler l'accès.
- **Réutilisation future**. Il se peut que d'autres personnes ne puissent pas utiliser le matériau sans une documentation adéquate.
- **Coût**. Il est coûteux de créer des métadonnées manuellement et il n'est pas toujours facile de générer automatiquement des métadonnées de préservation. Des métadonnées supplémentaires pour les besoins de la préservation numérique nécessitent donc un compromis coût/bénéfice prudent.

Standard PREMIS (PREservation Metadata : Implementation Strategies)

PREMIS (*PREservation Metadata : Implementation Strategies*) est le standard international pour les métadonnées destinées à faciliter la préservation des objets numériques et à garantir leur utilisation à long terme. Développé par une équipe internationale, PREMIS est mis en œuvre dans des projets de préservation numérique dans le monde entier, et le soutien à PREMIS est intégré dans un certain nombre d'outils et de systèmes de préservation numérique commerciaux et open source.

Le dictionnaire de données PREMIS (PREMIS, 2013) est organisé autour d'un modèle de données composé de cinq entités associées au processus de préservation numérique :



1. **Entité intellectuelle** : un ensemble cohérent de contenus qui est décrit comme une unité : par exemple, un livre
2. **Objet :** une unité d'information distincte sous forme numérique, par exemple un fichier PDF
3. **Événement** : une action de préservation, par exemple l'entrée du fichier PDF dans l'entrepôt numérique
4. **Agent** : personne, organisation ou programme logiciel associé à un événement, par exemple l'éditeur d'un fichier PDF
5. **Droits** : une ou plusieurs autorisations relatives à un Objet, par exemple, l'autorisation de faire des copies du fichier PDF à des fins de préservation

Prises ensemble, les unités sémantiques définies dans le dictionnaire de données PREMIS représentent les informations essentielles nécessaires pour mettre en œuvre les activités de préservation numérique dans la plupart des contextes des entrepôts numériques. Toutefois, le concept de noyau dans le cadre de PREMIS est défini de manière vague : toutes les unités sémantiques ne sont pas considérées comme obligatoires dans toutes les situations, et certaines sont facultatives dans toutes les situations. Le dictionnaire de données tente de trouver un équilibre entre le fait de reconnaître qu'il y aura un chevauchement important des exigences en matière de métadonnées dans différents contextes d'entrepôts numériques, tout en reconnaissant que tous les contextes sont différents d'une manière ou d'une autre et que, par conséquent, leurs exigences respectives en matière de métadonnées seront rarement exactement les mêmes.

Mise en œuvre

Bien que le dictionnaire de données PREMIS ne soit pas une norme officielle, dans la mesure où il est géré par un organisme de normalisation reconnu, il a atteint le statut de norme acceptée pour les métadonnées de préservation dans la communauté de la préservation numérique. L'un des points forts mais aussi l'une des limites du dictionnaire de données PREMIS est qu'il doit être adapté aux exigences du contexte spécifique ; il ne s'agit pas d'une solution clé en main dans le sens où un entrepôt numérique se contenterait de mettre en œuvre globalement le dictionnaire

de données. Seule une partie peut être pertinente dans certaines circonstances de préservation numérique ; d'autre part, l'entrepôt numérique peut trouver que des informations supplémentaires à celles définies dans le dictionnaire sont nécessaires pour répondre à ses besoins. Par exemple, le dictionnaire de données ne prévoit pas de documenter les informations sur les dépendances commerciales et politiques d'un entrepôt numérique, qui peuvent être nécessaires pour appuyer la prise de décision en matière de préservation.
En bref, chaque entrepôt devra faire un certain effort pour adapter les normes de métadonnées de préservation et de documentation à son contexte et exigences particulières.

Lors de la mise en œuvre, chaque institution identifie normalement ce qui constitue le noyau minimal d'information nécessaire aux éléments catalogués de la collection. Chaque institution peut également identifier ses niveaux préférés de métadonnées et de documentation pour les acquisitions et peut notifier et encourager les fournisseurs ou les déposants à fournir ces informations. Le personnel examine et révise les informations fournies pour s'assurer qu'elles sont conformes aux directives de l'institution et il génère des notices de catalogue pour les données déposées en incorporant des normes de catalogage et de documentation afin de garantir que les informations sur ces objets peuvent être mises à la disposition des utilisateurs par le biais de catalogues appropriés. Dans de nombreux cas, les informations contextuelles relatives aux ressources seront cruciales pour leur utilisation future et cet aspect de la documentation ne doit pas être négligé.

Le niveau de catalogage et de documentation accompagnant un objet ou ajouté ultérieurement à celui-ci, ainsi que les limites qu'il peut imposer, peuvent être documentés au profit des futurs utilisateurs. Lorsque les ressources des données sont gérées par des tiers mais mises à disposition par l'intermédiaire d'une institution, les informations peuvent être fournies par le tiers sous une forme convenue qui est conforme aux directives de l'institution ou dans le format natif du fournisseur.

Lorsqu'un besoin d'accès renforcé existe, une institution peut s'engager à améliorer la documentation et le catalogage des informations à un niveau plus élevé pour répondre aux nouvelles exigences. Une amélioration rétrospective de la documentation ou du catalogue doit également avoir lieu lorsque la validation ou l'audit de la documentation et du catalogage d'une ressource montre que celle-ci est inférieure à une norme minimale acceptable.

Un nombre important d'utilisateurs et de fournisseurs de métadonnées de préservation ont adopté PREMIS et ont ainsi pu surmonter nombre des obstacles initiaux à sa mise en œuvre. Le processus de mise en œuvre de PREMIS dans un environnement de travail est facilité par un certain nombre d'outils qui peuvent extraire les métadonnées des objets numériques et produire des fichiers PREMIS XML. La *PREMIS Maintenance Activity* tient à jour une page web qui énumère les principaux outils disponibles pour l'utilisation de PREMIS. Elle comprend également une liste de discussion active via courrier électronique et un wiki pour le partage de documents. Pour plus d'informations, voir la section Ressources et Études de cas ci-dessous.
Voir aussi les sections connexes du manuel, notamment Acquisition et évaluation, et Planification de la préservation.

Ressources

Dictionnaire de données PREMIS pour les métadonnées de préservation, version 3.0
http://www.loc.gov/standards/premis/v3/index.html
Le dictionnaire de données PREMIS et sa documentation d'appui constituent une ressource complète et pratique pour la mise en œuvre des métadonnées de préservation dans les entrepôts numériques. Le dictionnaire de données est construit sur un modèle de données qui définit cinq entités : Entités intellectuelles, objets, événements, droits et agents. Chaque unité sémantique définie dans le dictionnaire de données est une propriété de l'une des entités du modèle de données. La version 3.0 a été publiée en juin 2015 et mise à jour en novembre 2015 (273 pages).

Introduction en français : Priscilla Caplan, *Comprendre PREMIS, Library of CongressNetwork Development and MARC Standards Office*, 2009. Traduction par Louise Fauduet, Clément Oury, Sébastien Peyrard, Jean-Philippe Tramoni, Bibliothèque nationale de France.
https://www.loc.gov/standards/premis/Understanding-PREMIS_french.pdf

Métadonnées de préservation (2e édition), Rapport de veille technologique de la DPC
http://dx.doi.org/10.7207/twr13-03
Ce rapport se concentre sur les nouveaux développements dans le domaine des métadonnées de préservation rendus possibles par l'émergence de PREMIS comme standard international de facto. Il se concentre sur les principaux sujets de mise en œuvre, notamment les révisions du dictionnaire de données, la sensibilisation de la communauté, le conditionnement (en mettant l'accent sur le METS), les outils, les mises en œuvre de PREMIS dans les systèmes de préservation numérique et les ressources de mise en œuvre. Publié en 2013 (36 pages).

Outils pour la mise en œuvre des métadonnées de préservation
http://www.loc.gov/standards/premis/tools_for_premis.php
La *PREMIS Maintenance Activity* tient à jour une page web qui énumère les principaux outils disponibles pour utiliser PREMIS. Cette page contient des entrées sur les outils, en plus de pointeurs vers d'autres outils qui peuvent être utilisés pour générer des fichiers METS (*Metadata Encoding and Transmission Standard* - un schéma XML pour le conditionnement des métadonnées des objets numériques) en conjonction avec PREMIS. La majorité des outils répertoriés servent à extraire les métadonnées techniques des objets numériques et à les convertir pour les encoder dans l'entité PREMIS Object. D'autres outils peuvent être utilisés pour vérifier les formats, ou valider les fichiers par rapport à leurs empreintes.

Site web PREMIS
http://www.loc.gov/standards/premis/index.html
Le comité éditorial PREMIS coordonne les révisions et la mise en œuvre du standard PREMIS, qui comprend le dictionnaire de données, un schéma XML et la documentation d'appui. Le forum du groupe des utilisateurs PREMIS, hébergé par la *PREMIS Maintenance Activity*, comprend une liste de discussion active via courrier électronique et un wiki pour le partage des documents. Le wiki est une ressource particulièrement utile pour les nouveaux utilisateurs, car il comprend des documents issus des tutoriels PREMIS, une collection d'exemples d'utilisation de PREMIS et des liens vers des informations sur les outils PREMIS. La *PREMIS Maintenance Activity* tient un registre actif des mises en œuvre de PREMIS.

Documentation de vos données
http://www.data-archive.ac.uk/create-manage/document[43]
Un excellent ensemble de ressources pour aider les chercheurs dans la documentation et les métadonnées de leurs études de recherche, rassemblées par la *UK Data Archive*.

Lignes directrices du service de données archéologiques à l'intention des déposants
http://archaeologydataservice.ac.uk/advice/guidelinesForDepositors
Les lignes directrices du SDA à l'intention des déposants fournissent des conseils sur la manière de préparer correctement les données et de compiler les métadonnées en vue de leur dépôt auprès du SDA et décrivent les modalités de dépôt des données. Il existe également une série de fiches de travail et de listes de contrôle succinctes couvrant : la gestion des données ; la sélection et la durée de préservation ; les formats de fichier et les métadonnées préférés. D'autres ressources à l'usage des déposants potentiels comprennent une série de guides de bonnes pratiques, qui complètent les lignes directrices du SDA et fournissent des informations plus détaillées sur des types de données spécifiques.

Études de cas

Note de cas de la DPC : *British Library* ASR2 utilisant le METS pour garder les données et les métadonnées ensemble pour la préservation
http://www.dpconline.org/component/docman/doc_download/474-casenoteasr2.pdf
Cette étude de cas financée par le Jisc examine le projet *Archival Sound Recordings* 2 de la *British Library*, en notant que l'un des défis pour l'accès à long terme au contenu numérisé est de s'assurer que les informations descriptives et le contenu numérisé ne sont pas séparés les uns des autres. La *British Library* a utilisé un standard appelé METS pour éviter cela. Juillet 2010 (4 pages).

[43]NdT : Le document n'était plus accessible en ligne à la date de la présente traduction, une archive est disponible sur :https://web.archive.org/web/20121106190403/http://www.data-archive.ac.uk/create-manage/document

Gunia B., Sandusky J., “Designing metadata for long-term data preservation: DateONE case study”, 2011
https://doi.org/10.1002/meet.14504701435
Une brève description de la façon dont PREMIS a été utilisé pour spécifier les besoins de préservation des métadonnées pour les données scientifiques de DataONE (*Data Observation Network for Earth*), 2010 (2 pages).

Étude de cas Preservica : Questions-réponses avec Glen McAninch, Kentucky Department for Libraries and Archives
https://preservica.com/uploads/resources/Preservica-Kentucky-QA-2014_NEW.pdf
Glen McAninch discute de l'importance de la provenance, du contexte et des métadonnées dans la préservation des documents d'archives numériques.

Registre des implémentations PREMIS
http://www.loc.gov/standards/premis/registry/index.php
La *PREMIS Maintenance Activity* tient un registre actif de plus de 40 mises en œuvre de PREMIS avec des détails sur l'entrepôt numérique et son utilisation de PREMIS. Bien qu'il ne s'agisse pas officiellement d'études de cas, les enregistrements contiennent des détails sur l'expérience pratique, par exemple la création d'un entrepôt numérique aux Archives nationales suédoises à l'aide de PREMIS.

Références

Gartner, R. and Lavoie, B., 2013. *Preservation Metadata (2nd edition),DPC Technology Watch Report* 13-3 May 2013. Disponible à l'adresse suivante : http://dx.doi.org/10.7207/twr13-03
PREMIS, 2015.*Data Dictionary for Preservation Metadata*, Version 3.0. Disponible à l'adresse suivante : http://www.loc.gov/standards/premis/v3/index.html

Solutions et outils techniques

Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

A qui s'adresse-t-il ?

Les encadrants (*DigCurV : Manager Lens*) et le personnel opérationnel (*DigCurV : Practitioner Lens*), les éditeurs et autres créateurs de données, les fournisseurs de services tiers.

Niveau de connaissance présumé

De novice à intermédiaire.

Objectif

- Se concentrer sur les outils techniques et les applications qui soutiennent la préservation numérique : logiciels, applications, programmes et services techniques.
- Envisager le déploiement pratique de techniques et de technologies de préservation, qu'il s'agisse de programmes relativement petits et discrets (comme DROID) ou de solutions à l'échelle de l'entreprise qui intègrent de nombreux outils.
- Cette section exclut d'autres questions et normes plus stratégiques ou politiques qui sont parfois décrites comme des outils : elles sont traitées ailleurs dans le manuel.

Outils

Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Un guide d'initiation aux outils de préservation numérique

L'utilité des outils techniques pour la préservation numérique dépend du contexte dans lequel ils sont déployés. Une recommandation communautaire peut être forte, mais si elle ne correspond pas à votre fonction spécifique ou à votre contexte organisationnel, alors il y a de fortes chances que l'outil ne remplisse pas son rôle. Avant de choisir des outils de préservation numérique, il est donc important d'examiner attentivement le processus de travail technique et le cadre institutionnel dans lequel ils s'inscrivent. Un exemple pratique a été présenté par *Northumberland Estates* qui a développé un cadre d'évaluation simple pour évaluer les outils dans leur contexte.
Une autre façon d'aborder ce sujet consiste à examiner dans quelle mesure un outil donné permettra de mener des actions de préservation découlant d'un plan de préservation convenu, qui à son tour découle d'un cadre politique donné.

Réflexion sur les outils de préservation numérique

Les problèmes suivants sont fréquemment rencontrés dans le processus de déploiement des outils de préservation numérique. Cette liste n'est pas exhaustive, mais l'examen de ces questions permettra de faire des choix judicieux et réalistes.

Logiciels Open Source et logiciels commerciaux

Certaines organisations - souvent dans l'enseignement supérieur et surtout dans les dépôts d'institutions de recherche - sont à l'aise avec l'utilisation de logiciels Open Source, surtout lorsqu'elles disposent d'un groupe de développeurs en interne. Le logiciel « *Open Source* » est un logiciel dont le code sous-jacent est mis à disposition gratuitement, offrant la liberté d'ajouter, de modifier ou de développer celui-ci. D'autres organisations qui n'ont pas facilement accès à des développeurs ont tendance à avoir des règles de passation de marché qui préfèrent des solutions commerciales « sur étagère » soutenues par des contrats de maintenance. La distinction entre les logiciels à code source ouvert et les logiciels commerciaux est souvent exagérée car les deux s'influencent mutuellement. Néanmoins, vous devrez peut-être tenir compte des normes et de la culture de votre organisation lorsque vous choisirez des outils.

Solutions au niveau de l'entreprise et micro-services

Certains outils de préservation numérique sont conçus pour offrir des solutions « du sol au plafond », c'est-à-dire qu'ils fournissent un processus intégré de bout en bout qui permet de

fournir toutes (ou quasiment toutes) les fonctions de préservation numérique à une organisation. En fait, les solutions d'entreprise sont le plus souvent construites de telle manière qu'elles regroupent des outils individuels intégrés dans une interface unique. La solution à un problème donné peut être relativement simple et votre organisation peut se satisfaire de cet assemblage d'une série de petits outils pour des fonctions distinctes. Cette option peut permettre des progrès rapides et s'avère utile pour tester et essayer les outils mais peut être difficile à maintenir sur une longue période. Dans d'autres organisations, le contrôle du déploiement des logiciels est beaucoup plus strict et l'on s'attend à ce que les solutions soient construites sur l'ensemble d'un processus de travail - ce qui exige des solutions complètes. Cette façon de faire peut être moins rapide mais plus facile à maintenir sur le long terme. Avant de choisir un outil, il convient donc de prendre en compte la position de votre organisation sur ce spectre.

Description des processus de travail

L'un des éléments clés à prendre en compte pour les outils est leur place dans un processus de travail global. Ainsi, avant de choisir des outils, il est utile d'examiner et de définir l'ensemble du processus de travail. Le fait d'avoir explicité un processus de travail peut également aider à identifier les processus redondants ainsi que les principaux goulets d'étranglement. Fréquemment, des outils résolvent un problème dans une partie d'un processus de travail, pour ensuite créer un problème ailleurs. En outre, les organisations peuvent avoir plusieurs processus de travail avec des exigences différentes qui entrent en conflit d'une manière ou d'une autre. La description d'un processus de travail fournit donc une base pour anticiper les difficultés et peut fournir une feuille de route pour une amélioration continue.

Spécifications d'exigences claires

Afin d'évaluer l'utilité et la valeur pour votre organisation des nombreux outils disponibles, il est utile de disposer d'une déclaration explicite des exigences. Les outils peuvent être comparés et étalonnés de manière transparente et les décisions justifiées en conséquence. Bien exécutées, les activités de collecte des exigences peuvent impliquer toute une série de parties prenantes et donc maximiser le potentiel d'alignement et d'efficacité, en atteignant des objectifs stratégiques et organisationnels plus larges.

Exigences changeantes et évolutives

Il est normal que les exigences changent au fil du temps. En effet, la préservation numérique vise essentiellement à relever les défis liés aux changements technologiques inévitables. Il est donc nécessaire de contrôler et de revoir les outils pour s'assurer qu'ils restent adaptés à leurs objectifs et que toute modification des exigences est rendue explicite. Un examen périodique de la spécification des exigences est recommandé.

Pérennité des outils et participation de la communauté

Le critère de pérennité est un facteur important dans toute décision concernant les outils que vous utilisez pour la préservation numérique. La pérennité en termes d'outils peut inclure une base d'utilisateurs actifs, une maintenance et un développement. Par exemple, une base d'utilisateurs importante, à la fois en termes de fournisseurs commerciaux et de logiciels libres, peut être un indicateur vital pour identifier un outil viable. Il convient de noter qu'une communauté peut changer rapidement et pour des raisons qui ne sont pas facilement prévisibles. Les « derniers entrants dans le domaine » peuvent rapidement devenir des leaders tandis que de grandes communautés peuvent se réduire aussi vite que les nouvelles technologies prennent le dessus sur

les technologies existantes. Par conséquent, il peut être nécessaire de surveiller la bonne santé de la communauté des développeurs qui maintiennent vos outils.

Trouver des outils de préservation numérique : outils et registres d'outils

L'une des caractéristiques les plus appréciées de la préservation numérique au cours des deux dernières décennies a été le développement rapide de logiciels, d'outils et de services qui améliorent et facilitent les processus de préservation numérique. Au fur et à mesure que la communauté de la préservation numérique a grandi en taille et en sophistication, nos outils sont devenus plus puissants et plus perfectionnés. Cette prolifération et cette spécialisation accrue peuvent également constituer un obstacle au déploiement, en particulier lorsque les outils sont le fruit de projets de recherche relativement courts et de portée limitée. Par conséquent, la diversité des outils peut sembler de plus en plus déconcertante pour les nouveaux utilisateurs, tandis que l'accès au marché pour les développeurs est de plus en plus compliqué.

Les registres d'outils sont apparus ces dernières années comme un moyen d'aider les utilisateurs à trouver les outils dont ils ont besoin. Il existe aujourd'hui un certain nombre de registres qui décrivent les outils de préservation numérique. En fonction des intérêts des personnes qui les utilisent, ils peuvent également fournir des descriptions détaillées, des critiques ou des commentaires sur les outils de la communauté au sens large. Ils ne sont donc pas seulement utiles aux utilisateurs : en permettant aux experts d'examiner les outils et d'évaluer leurs performances, ils signalent les forces et les faiblesses et fournissent une base pour les développements futurs ; en mettant les outils en relation avec les utilisateurs, ils aident les développeurs à atteindre un public beaucoup plus large et à obtenir des commentaires pour améliorer leurs outils.

Les registres sont un moyen classique pour la communauté de la préservation numérique de partager l'information. Il existe d'autres types de registres, tels que les « registres de formats » qui décrivent les performances de formats de fichiers donnés, ou les « registres d'environnement » qui décrivent l'empilement des technologies nécessaire pour créer un environnement d'exécution permettant d'émuler ou de virtualiser des logiciels. Ces questions sont traitées ailleurs dans le manuel.

Y a-t-il trop de registres ?

Si les registres sont un bon moyen de gérer la prolifération des outils, il est maintenant reconnu que la prolifération des registres est également un obstacle potentiel à leur utilisation. Le registre COPTR a été conçu spécifiquement pour résoudre ce problème, en s'appuyant sur des données provenant de multiples sources, dont le DCC, le POWRR et la Bibliothèque du Congrès.

Soutien et conseils pratiques

Après avoir examiné certains des registres d'outils et des outils de préservation numérique qui sont à la disposition des organisations, la question qui se pose souvent est de savoir lequel choisir en fonction de l'objectif de votre organisation. Avant tout, il est important que votre choix soit aligné sur les besoins et l'orientation stratégique de votre organisation ; les ressources et les études de cas ci-dessous fournissent des outils d'évaluation et des conseils pour soutenir une mise en œuvre réussie.

Ressources

Registre d'outils

« *Community Owned digital Preservation Tool Registry* » COPTR

http://coptr.digipres.org/Main_Page

Le registre COPTR décrit des outils utiles pour la préservation numérique à long terme et agit principalement comme un outil de recherche et d'évaluation pour aider les professionnels à trouver les outils dont ils ont besoin pour préserver les données numériques. Le registre COPTR vise à rassembler en un seul endroit les connaissances de la communauté de la préservation numérique sur les outils de préservation. Il a été initialement alimenté par des données provenant de registres gérés par les organisations partenaires du registre COPTR, notamment ceux tenus par le *Digital Curation Centre*, le *Digital Curation Exchange*, la *National Digital Stewardship Alliance*, l'*Open Preservation Foundation* et le *Preserving digital Objects With Restricted Resources project* (POWRR). Le registre COPTR permet de saisir des détails factuels de base sur un outil, ses fonctions, la manière de trouver plus d'informations (URL pertinentes) et des références aux expériences des utilisateurs avec l'outil. Le champ d'application est une interprétation large du terme « préservation numérique ». En d'autres termes, si un outil est utile pour effectuer une fonction de préservation numérique telle que celles décrites dans le modèle OAIS ou le modèle de cycle de vie DCC, alors il entre dans le champ d'application de ce registre.

Liste des outils *AV Preserve*

https://www.weareavp.com/products/

Une liste d'outils particulièrement utiles pour la préservation à long terme des objets audiovisuels, qu'ils soient numérisés ou nativement numériques.

Liste des outils et services du DCC

https://www.dcc.ac.uk/tools

Le DCC est un centre d'excellence qui soutient les chercheurs du Royaume-Uni dans leurs efforts pour relever les défis de la préservation et de la préservation des ressources numériques. Pour atteindre cet objectif, il a offert un certain nombre de services de maintenance et de conseil soutenus par une recherche et un développement ciblés. Le premier comprend un catalogue d'outils et de services qui catégorise les outils destinés aux chercheurs et aux conservateurs. Les informations sont également intégrées dans COPTR (voir ci-dessus).

Registre du DCH-RP

https://www.dchrp.eu/index.php?en/137/registry-of-services-tools

Le registre des outils de la « Feuille de route pour la préservation du patrimoine culturel numérique » (DCH-RP) a recueilli et décrit les informations et les connaissances relatives aux outils, technologies et systèmes pouvant être appliqués à des fins de préservation du patrimoine culturel numérique. La version 3 du registre a été créée en 2014.

Inventaire des FLOSS (*Free/libre open-source software*) dans le domaine du patrimoine culturel

https://docs.google.com/spreadsheets/d/1bOoQiXFjGyR3oEubdLdkfCat7V4TsNLnEXGOJWkJ63c/edit#gid=516255520&vpid=D2

Produit par le projet Europeana, financé par l'Union Européenne, cet inventaire répertorie les logiciels libres et gratuits qui peuvent être utiles dans le secteur du patrimoine culturel. Bien qu'il

ne se limite pas aux outils de préservation numérique, l'inventaire contient des informations sur une variété d'outils ayant des applications dans la préservation numérique, en évaluant leur objectif, la qualité de la documentation, le niveau de soutien, les exigences en matière de licence et en fournissant des liens vers des informations sur le projet et le code source.

Vitrine des outils NDIIPP de la Bibliothèque du Congrès
http://www.digitalpreservation.gov/tools/
Le registre des outils de préservation numérique de la Bibliothèque du Congrès est une liste sélective d'outils et de services intéressant les personnes travaillant dans le domaine de la préservation numérique. Il n'est plus activement tenu à jour et le contenu est intégré dans COPTR (voir ci-dessus).

Grille d'outils pour la préservation des objets numériques avec des ressources restreintes (POWRR)
http://digitalpowrr.niu.edu/tool-grid/
POWRR a étudié, évalué et recommandé des solutions de préservation numérique évolutives et durables pour les organisations ayant des quantités de données relativement faibles et/ou moins de ressources. Un résultat important du projet a été la grille d'outils produite au début de 2013, basée sur les catégories fonctionnelles du modèle de référence OAIS. Une version actualisée de la grille d'outils POWRR peut désormais être générée dans COPTR (voir ci-dessus).

Questions et réponses sur la préservation numérique
http://qanda.digipres.org/
Il s'agit d'un site où vous pouvez poster des questions et des réponses afin de vous aider mutuellement à utiliser au mieux les outils, techniques, processus, flux de travail, pratiques et approches visant à assurer l'accès à long terme à l'information numérique. Les questions-réponses sur la préservation numérique sont actuellement animées par des représentants des organisations membres de l'ASDN et de l'OPF.

Dossiers électroniques pratiques
http://e-records.chrisprom.com/author/prom/
Le blog *Software and Tools for Archivists* de Chris Prom. Bien que certaines informations datent de plusieurs années, le blog constitue un point de départ utile pour comprendre les utilisations de divers outils de préservation numérique et une évaluation standardisée des outils par rapport à des critères définis, notamment la facilité d'installation, la convivialité, l'extensibilité, etc. En plus des informations sur les outils, le blog contient une foule d'autres ressources utiles, notamment des modèles de politiques et de processus de travail, des approches recommandées.

Études de cas

Journal d'un projet de préservation d'un dépôt d'archives
http://blog.soton.ac.uk/keepit/
Un journal des avancements a été réalisé (entre avril 2009 et septembre 2010), alors que le projet KeepIt, financé par le Jisc, s'attaquaient aux défis de la préservation des contenus des dépôts

numériques dans la recherche, l'enseignement, les sciences et les arts. Il comprend une expérience utile pour évaluer les outils de préservation.

Northumberland Estates

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

Northumberland Estates a développé un cadre d'évaluation simple pour évaluer les outils dans leur contexte. Le projet visait à étudier les options de dépôt numérique actuellement disponibles pour les petites et moyennes organisations disposant de ressources limitées. Notez que les recommandations formulées dans la dernière analyse de rentabilité reflètent les besoins organisationnels de *Northumberland Estates* et peuvent ne pas correspondre à vos propres objectifs. L'étude de cas a été préparée dans le cadre du projet SPRUCE, financé par le Jisc.

Intégrité et empreintes

Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Intégrité

« L'intégrité, au sens de la préservation, signifie l'assurance qu'un fichier numérique est resté inchangé, autrement dit figé. » (Bailey, 2014). L'intégrité ne s'applique pas seulement aux fichiers, mais à tout objet numérique qui contient une série de bits où ce « flux d'octets » doit être maintenu intact si l'on a la connaissance que rien n'a changé. L'intégrité peut s'appliquer aux images ou aux vidéos contenues dans un objet audiovisuel, aux fichiers individuels dans un zip, aux métadonnées dans une structure XML, aux enregistrements dans une base de données ou aux objets dans un magasin d'objets. Cependant, les fichiers sont actuellement le moyen le plus courant de stocker des objets numériques et l'intégrité des fichiers peut être établie et contrôlée grâce à l'utilisation d'empreintes numériques.

Empreintes

L'empreinte d'un fichier est « une empreinte numérique », c'est-à-dire que même la plus petite modification apportée au fichier entraîne une modification complète de l'empreinte. Les empreintes sont généralement créées à l'aide de techniques cryptographiques et peuvent être générées à l'aide d'une série d'outils facilement accessibles et de logiciels libres. Il est important de noter que si les empreintes peuvent être utilisées pour détecter si le contenu d'un fichier a changé, elles ne vous indiquent pas où a eu lieu le changement dans le fichier.
Les empreintes ont trois rôles principaux :

1. Savoir qu'un fichier a été correctement reçu d'un propriétaire ou d'une source de contenu, puis transféré avec succès au stockage de préservation.
2. Savoir que l'intégrité du fichier a été maintenue lorsque ce fichier est stocké.
3. Leur diffusion aux utilisateurs du fichier à l'avenir afin qu'ils puissent savoir que le fichier a été correctement récupéré depuis le lieu de stockage et leur a été livré.

Cela permet d'établir une « chaîne de confiance » entre ceux qui produisent ou fournissent les objets numériques, ceux qui sont responsables de leur stockage permanent et ceux qui doivent utiliser l'objet numérique qui a été stocké. Dans le modèle de référence OAIS (ISO, 2012), ces personnes sont les producteurs, l'OAIS lui-même – le dépositaire – et les utilisateurs.

Application dans le domaine de la préservation numérique

Si une organisation possède plusieurs copies de ses fichiers, par exemple comme recommandé dans la section Stockage, les empreintes peuvent être utilisées pour contrôler l'intégrité de chaque

copie d'un fichier et si l'une des copies a changé, alors une des autres copies peut être utilisée pour créer une nouvelle copie de remplacement correcte. La méthode consiste à calculer régulièrement une nouvelle empreinte pour chaque copie d'un fichier et à la comparer avec la valeur de référence dont on sait qu'elle est correcte. Si un écart est constaté, on sait que le fichier a été corrompu d'une manière ou d'une autre et qu'il devra être remplacé par une nouvelle copie non corrompue. Ce processus est connu sous le nom de « *Data scrubbing* » (nettoyage des données).

Les empreintes sont idéales pour détecter si des modifications indésirables ont été apportées aux documents numériques. Cependant, il arrive que les objets numériques soient modifiés délibérément, par exemple si un format de fichier est migré. Cela entraîne une modification de l'empreinte. Il faut alors établir de nouvelles empreintes après la migration, qui deviennent le moyen de vérifier l'intégrité des données du nouveau fichier à l'avenir.
Les fichiers doivent être régulièrement comparés à leurs empreintes. La fréquence des contrôles dépend de nombreux facteurs, notamment du type de stockage, de la qualité de sa maintenance et de la fréquence de son utilisation. En règle générale, la vérification des bandes de données peut être effectuée chaque année et celle des systèmes sur disque dur tous les six mois. Des contrôles plus fréquents permettent de détecter et de résoudre les problèmes plus rapidement, mais au prix d'une charge plus importante sur le système de stockage et de ressources de traitement plus importantes.

Les empreintes peuvent être stockées de différentes manières, par exemple dans un enregistrement PREMIS, dans une base de données ou dans un « manifeste » qui accompagne les fichiers dans un système de stockage.

Les outils pour la génération et l'utilisation des empreintes bénéficient d'un bon support. Comme il s'agit de fonctions relativement simples, les empreintes sont intégrées dans de nombreux autres outils de préservation numérique. Par exemple, la génération d'empreintes dans le cadre du processus de versement et l'ajout de ces informations d'intégrité aux Paquets d'Information Archivé (AIP) générés, ou la possibilité de générer des manifestes d'empreintes pour plusieurs fichiers et de regrouper le manifeste et les fichiers pour faciliter le transport ou le stockage. En outre, md5sum et md5deep fournissent des outils simples en ligne de commande qui fonctionnent sur toutes les plateformes pour générer des empreintes sur des fichiers ou des répertoires individuels.
Il existe plusieurs algorithmes d'empreintes différents, par exemple MD5 et SHA-256, qui peuvent être utilisés pour générer des empreintes de plus en plus fortes. Plus l'algorithme est « fort », plus il est difficile de modifier délibérément un fichier de manière non détectable. Cela peut être important pour les applications où il est nécessaire de démontrer la robustesse face à la corruption ou à l'altération malveillante d'objets numériques, par exemple lorsque la valeur probante et la recevabilité juridique sont importantes. Toutefois, si des empreintes sont utilisées pour détecter la perte ou l'endommagement accidentel de fichiers, par exemple en raison d'une défaillance de stockage, alors le MD5 est suffisant et présente l'avantage d'être bien pris en charge par les outils et d'être rapide à calculer.

Le manuel suit les niveaux de préservation de la *National Digital Stewardship Alliance* (NDSA) (NDSA, 2013) en recommandant quatre niveaux auxquels la préservation numérique peut être prise en charge par des techniques de calcul d'empreintes des fichiers et d'intégrité des données. Les principaux bénéfices de la vérification de l'intégrité ne peuvent être atteints que s'il existe

plusieurs copies des objets numériques, par exemple en permettant la réparation si l'intégrité de l'une des copies a été perdue.

Niveau	Activité	Risques traités et avantages obtenus
1	• Vérifier l'intégrité du versement au niveau du fichier s'il a été fourni avec le contenu. • Créer une information d'intégrité si elle n'a pas été fournie avec le contenu.	• Les objets numériques corrompus ou incorrects ne sont pas sciemment stockés. • L'authenticité des objets numériques peut être affirmée. • L'intégrité de base est établie de sorte que les modifications de données non souhaitées puissent être détectées.
2	• Vérifier l'intégrité de tous les versements. • Utiliser des bloqueurs d'écriture lorsque vous travaillez avec des supports originaux. • Contrôler les virus sur les contenus à haut risque.	• Aucun objet numérique dont l'intégrité n'est pas confirmée ne peut entrer dans le stockage de préservation. La valeur probante est prise en charge pour garantir l'authenticité. • L'assurance peut être donnée à tous les fournisseurs de contenu que leur contenu a été reçu en toute sécurité. Les supports originaux sont protégés. • Aucun contenu malveillant ne peut entrer dans le stockage de préservation.
3	• Vérifier à intervalles réguliers l'intégrité du contenu conservé sur les systèmes de stockage de préservation. • Tenir à jour les journaux d'informations sur l'intégrité et fournit un audit à la demande. • Pouvoir détecter les données corrompues. • Contrôler les virus sur l'ensemble du contenu.	• Protection contre un large éventail de cas de corruption et de perte de données. Les problèmes de stockage sont détectés plus tôt. • La corruption ou la perte de données ne passent pas inaperçues en raison d'« erreurs silencieuses » ou de « défaillances non détectées ». Les objets numériques ne sont pas dans un état d'intégrité « inconnu ». • Il est possible d'apporter en permanence la preuve que les documents numériques sont intacts et corrects.
4	• Vérifier l'intégrité de tout le contenu en réponse à des événements ou activités spécifiques • Pouvoir remplacer/réparer les données corrompues. • Veiller à ce qu'aucune personne n'ait accès en écriture à toutes les copies.	• Les cas de défaillance qui menacent les objets numériques sont contrés de manière proactive. Toutes les copies des documents numériques sont activement conservées. • Assurance aux utilisateurs de l'intégrité et de l'authenticité des objets numériques auxquels ils ont accès. • L'efficacité de l'approche de préservation peut être mesurée et démontrée. • Conformité aux normes, par exemple ISO 16363 Audit et certification des dépôts numériques fiables.

Blocage de l'écriture

Notez que la *National Digital Stewardship Alliance* (NDSA) recommande l'utilisation de bloqueurs d'écriture de niveau 2. Ceci a pour but d'empêcher l'accès en écriture aux supports sur lesquels les objets numériques pourraient se trouver avant d'être copiés dans le système de stockage de préservation. Par exemple, si les objets numériques sont livrés à une organisation sur un disque dur ou une clé USB, un bloqueur d'écriture empêchera la suppression accidentelle de ces objets numériques lorsque le disque ou la clé est lu. Les objets numériques peuvent ne pas se trouver sur un support physique, par exemple sur un serveur de stockage existant ou par un transfert de réseau, par exemple un téléchargement ftp. Dans ces cas, les bloqueurs d'écriture ne s'appliqueraient pas et d'autres mesures seraient utilisées pour rendre les objets numériques « en lecture seule » sur la source et donc immuables avant la confirmation que les objets numériques ont été transférés avec succès vers le stockage de préservation. Les bloqueurs d'écriture n'existent pas non plus pour tous les types de supports. Si un bloqueur d'écriture est applicable, les coûts/compétences nécessaires pour les utiliser doivent être mis en regard avec le risque de dommage aux objets numériques originaux ou la nécessité d'avoir une authenticité rigoureuse des données. Par conséquent, certaines organisations peuvent considérer que l'utilisation de bloqueurs d'écriture est inutile ou qu'il s'agit d'une étape de niveau 3 ou 4.

Ressources

Bailey, J., 2014, Protect Your Data: File Fixity and Data Integrity, The Signal, Bibliothèque du Congrès.
http://blogs.loc.gov/thesignal/2014/04/protect-your-data-file-fixity-and-data-integrity/

Vérification de votre contenu numérique : Qu'est-ce que l'intégrité et quand devrais-je la vérifier ?
http://digitalpreservation.gov/ndsa/working_groups/documents/NDSA-Fixity-Guidance-Report-final100214.pdf?loclr=blogsig

De nombreux membres de la communauté des conservateurs savent qu'ils doivent vérifier l'intégrité de leur contenu, mais comment, quand et à quelle fréquence ? Ce document publié par la NDSA en 2014 vise à aider les responsables de la préservation à répondre à ces questions d'une manière qui soit logique pour leur organisation en fonction de leurs besoins et de leurs ressources (7 pages).

Outil d'intégrité d'AVPreserve
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
https://csrc.nist.gov/csrc/media/projects/cryptographic-standards-and-guidelines/documents/examples/sha256.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
Un bon aperçu qui ne se limite pas à l'histoire orale, cette vidéo fournit une brève introduction au rôle de l'empreinte dans la préservation numérique. Elle met en scène Doug Boyd, directeur du Centre Louie B. Nunn pour l'histoire orale des bibliothèques de l'Université du Kentucky (3 minutes 25 secondes).

Références

Bailey, J., 2014. Protect Your Data: File Fixity and Data Integrity. *The Signal*. [blog]. Disponible à l'adresse suivante : http://blogs.loc.gov/thesignal/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. Disponible à l'adresse suivante :https://www.iso.org/standard/57284.html

Version française : Modèle de référence pour un système ouvert d'archivage d'information (OAIS), Pratique recommandée, CCSDS 650.0-M-2 (F), Livre Magenta (juin 2012, version française d'octobre 2017)
https://public.ccsds.org/Pubs/650x0m2%28F%29.pdf

NDSA, 2013.*The NDSA Levels of Digital Preservation: An Explanation and Uses, version 1 2013*. National Digital Stewardship Alliance. Disponible à l'adresse suivante : http://www.digitalpreservation.gov/ndsa/working_groups/documents/NDSA_Levels_Archiving_2013.pdf

Traduction française de la v2 : Thomas Ledoux, Marion Humbert, Edouard Vasseur. Utilisation des Niveaux de Préservation Numérique : aperçu de la V2.0, traduction française. 2020 https://hal-bnf.archives-ouvertes.fr/hal-02551807/document

Formats de fichiers et normes

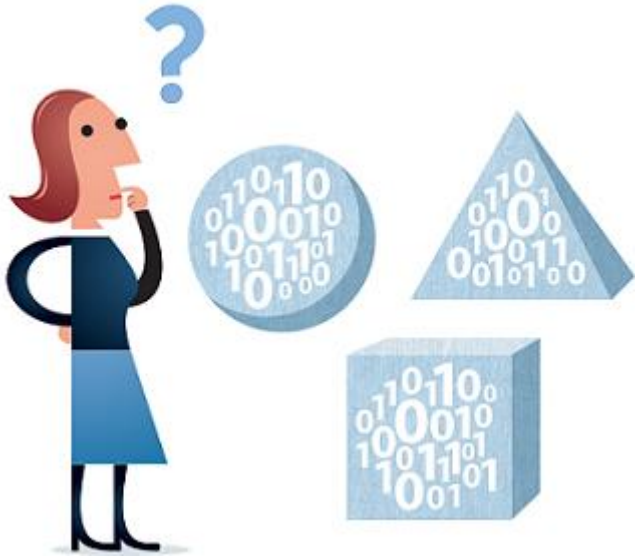


Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Introduction

La gestion des formats de fichiers doit être envisagée dans le contexte stratégique plus large de la planification de la préservation. Que peut se permettre votre organisation ? Quel effort de développement faudra-t-il fournir ? Qu'attendent les utilisateurs de vos collections ? Vous engagez-vous dans un problème de stockage ? À tout moment, la réponse aux questions de préservation numérique n'est pas d'essayer et de « tout faire ». Votre stratégie doit vous orienter vers des actions simples et pratiques, plutôt que d'essayer de prendre en charge plus de formats de fichiers que nécessaire.

L'objectif de cette section n'est pas de fournir une liste détaillée ou exhaustive des formats actuels pour différents types de contenus, mais d'attirer l'attention sur les implications plus larges des formats de fichiers pour leur application, et les conséquences pour leur préservation.

Une partie importante de ce chapitre se réfère à la sélection possible d'un format de fichier à des fins de migration. Si la migration est une stratégie de préservation valable, et assez courante pour de nombreux formats de fichiers, elle n'est pas la seule approche ou solution. Le cas échéant, le chapitre fera référence à d'autres méthodes possibles pour la préservation.

Formats de fichiers organisés par type de contenu

Différents types de contenu ont, au fil du temps, développé leurs propres formats de fichiers, car ils s'efforcent d'intégrer des fonctionnalités spécifiques à leurs besoins. Les principaux types de contenu sont les images, la vidéo, l'audio et le texte ; cependant, un nombre croissant de formats sont structurés pour répondre aux exigences des nouveaux médias, comme par exemple les formats pour les modèles 3D ou l'archivage web.

Les formats de fichiers varient énormément en termes de complexité, certaines données étant encodées en plusieurs couches. Dans certains cas, les formats de fichiers concernés ne sont qu'une partie d'un ensemble plus vaste, qui comprend des logiciels, du matériel et même des environnements informatiques complets.

Pour plus de conseils sur la préservation de types spécifiques de contenu numérique et de formats de fichiers associés, consultez les études de cas sur la préservation de contenus spécifiques dans le manuel.

Formats de fichiers - de quoi devrions-nous nous inquiéter ?

Obsolescence

Les formats évoluent au fur et à mesure que les utilisateurs et les développeurs identifient et intègrent de nouvelles fonctionnalités. Les nouveaux formats, ou versions de formats, peuvent introduire l'obsolescence des formats de fichiers à mesure que les nouvelles générations de logiciels suppriment progressivement la prise en charge des anciens formats. Lorsque les logiciels ne prévoient pas de rétrocompatibilité avec les anciens formats de fichiers, les données peuvent devenir inutilisables. Les formats commerciaux aussi bien que les formats Open Source sont vulnérables à l'obsolescence : les vendeurs utilisent parfois l'obsolescence planifiée pour inciter les clients à passer à de nouveaux produits, tandis que les communautés de logiciels Open Source peuvent supprimer le support des anciens formats si ceux-ci ne sont plus généralement requis par la communauté. L'obsolescence peut également être accidentelle : les entreprises et les communautés Open Source peuvent échouer.
L'obsolescence des formats de fichiers est un risque qu'il faut appréhender. Cela dit, le problème peut ne pas être aussi grave que la communauté de la préservation numérique le percevait il y a une dizaine d'années. De nombreux formats de fichiers établis existent toujours, sont toujours pris en charge et toujours utilisables. Il est fort probable que la majorité des formats de fichiers que vous traitez seront communément compris et bien pris en charge.

Prolifération

Dans certains secteurs, on peut dire que la prolifération est plus problématique que l'obsolescence. Si les formats ne sont pas normalisés, une organisation peut se retrouver avec un grand nombre de formats de fichiers différents, et des versions de ces formats : par exemple, de nombreuses versions différentes de formats PDF, Word, images, etc. Dans les domaines qui développent rapidement des formats de données sur mesure, ce problème peut être exacerbé. Le suivi et la gestion de tous ces formats, ceux qui sont à risque et les outils qui peuvent être utilisés pour chacun d'entre eux, peuvent constituer un défi majeur.
Votre stratégie de préservation numérique doit s'efforcer d'atténuer les effets de l'obsolescence et de la prolifération. Des stratégies telles que la migration, l'émulation, la normalisation et une sélection minutieuse des formats de fichiers sont toutes valables et méritent d'être prises en considération, dans le contexte de vos collections et de votre organisation.

Différents aspects des formats de fichiers au regard de la préservation numérique

Sélection des formats cibles pour la préservation

Tous les formats numériques ne sont pas adaptés ou conçus pour l'archivage ou la préservation. Toute politique de préservation doit donc reconnaître les exigences du contenu de la collection et décider du format de fichier qui préserve le mieux ces qualités. Associer le contenu à un choix approprié de format de préservation ou de format d'accès et identifier ce qui est important dans le contenu sont des éléments cruciaux.
Nous vous suggérons ci-dessous quelques facteurs à prendre en compte dans le choix de vos formats de fichier préférés :

Format Open Source ou propriétaire ?

Les formats Open Source, tels que JPEG2000, sont très populaires en raison de leur nature non propriétaire et du sentiment d'appropriation que les parties prenantes peuvent développer suite à leur utilisation. Toutefois, le choix entre les formats Open Source et les formats propriétaires n'est pas si simple et doit être examiné de près. Les formats propriétaires, tels que TIFF, sont considérés

comme très robustes ; cependant, ces formats seront en fin de compte susceptibles de connaître des problèmes de mise à niveau et d'obsolescence si leur propriétaire cesse ses activités ou développe une nouvelle alternative. De même, les formats Open Source peuvent être considérés comme technologiquement neutres, ne dépendant pas de modèles commerciaux pour leur développement ; cependant, ils peuvent également être considérés comme vulnérables aux susceptibilités des communautés qui les soutiennent.
Bien que ces formats non propriétaires puissent être sélectionnés pour de nombreux types de ressources, ce n'est pas le cas de tous. Pour de nombreux nouveaux domaines et applications, par exemple les systèmes d'information géographique ou la réalité virtuelle, seuls des formats propriétaires sont disponibles. Dans de tels cas, un facteur crucial sera les formats d'exportation pris en charge pour permettre aux données de sortir de ces environnements propriétaires ou d'y entrer.

Documentation et normes

La disponibilité de la documentation, par exemple, les spécifications publiées, est un facteur important dans le choix d'un format de fichier. La documentation peut exister sous la forme de spécifications du vendeur, d'une norme internationale, ou peut être créée et maintenue dans le contexte d'une communauté d'utilisateurs. Recherchez une norme bien documentée et largement appliquée. Assurez-vous que la norme figure dans le registre des formats de fichiers PRONOM.

Adoption

Un format de fichier sur lequel s'appuie un grand groupe d'utilisateurs crée beaucoup plus d'options pour ses utilisateurs. Il est utile de garder à l'esprit les niveaux d'utilisation et de prise en charge des formats dans le monde entier, mais aussi de découvrir ce que font des organisations similaires à la vôtre et de partager les meilleures pratiques en matière de sélection des formats. L'adoption à grande échelle d'un format peut vous donner plus de confiance dans votre stratégie de préservation.

Sans perte ou avec pertes ?

Les formats avec pertes sont ceux dans lesquels les données sont compressées, ou éliminées, lors de la phase de codage. Le format MP3 est largement utilisé pour la distribution commerciale de fichiers musicaux sur le web, car le processus d'encodage avec pertes entraîne une réduction de la taille des fichiers.
Le format TIFF est un exemple de format d'image capable de prendre en charge des données sans perte. Il peut contenir une image à haute résolution. Le format JPEG est un exemple de format de fichier image avec pertes. Sa polyvalence et sa petite taille de fichier en font un choix approprié pour créer une copie d'accès d'une image de taille plus petite pour la transmission sur un réseau. Il ne serait pas approprié de stocker l'image JPEG à la fois comme format d'accès et comme format d'archivage en raison de la perte irrémédiable de données que cela impliquerait.
Une règle empirique pourrait être de choisir des formats sans perte pour la création et le stockage des « masters d'archives » ; les formats avec pertes ne devraient être utilisés qu'à des fins de livraison/accès et ne devraient pas être considérés comme des archives. Une telle règle est particulièrement adaptée à un projet de numérisation, notamment d'images fixes.

Prise en charge des métadonnées

Certains formats de fichier prennent en charge les métadonnées, ce qui signifie que certaines métadonnées peuvent être directement inscrites dans une instance de fichier (par exemple,

JPEG2000 prend en charge certains champs de métadonnées de droits). Cela peut avoir son importance, en fonction de votre approche dans la gestion des métadonnées.

Propriétés significatives des formats de fichiers

Il s'agit d'un domaine complexe. Selon un point de vue, les propriétés significatives sont « l'essence » du contenu des fichiers ; une stratégie qui va au cœur de « ce qu'il faut préserver ». Qu'est-ce que la communauté des utilisateurs attend du rendu ? Quels aspects de l'original essayez-vous de préserver ? Cette stratégie pourrait signifier que vous n'avez pas à vous engager à préserver tous les aspects d'un format de fichier, mais seulement ceux qui ont le plus de sens et de valeur pour l'utilisateur.
Les propriétés significatives peuvent également faire référence à une gamme très spécifique de métadonnées techniques qui doivent être présentes pour qu'un fichier soit rendu (par exemple, la largeur de l'image). Certains outils de migration peuvent supprimer ces métadonnées, ou les perdre par d'autres actions de préservation dans le dépôt. La stratégie de préservation doit empêcher cette perte de se produire. Il devient donc important d'identifier, d'extraire, de stocker et de préserver les propriétés significatives à un stade précoce du processus de préservation.

Mesures à prendre

De nombreuses mesures peuvent être prises pour prendre en charge les formats de fichiers dans vos entrepôts numériques, et il existe de nombreux outils pour vous aider dans cette tâche. Ils sont maintenant si nombreux que des registres d'outils de préservation numérique sont en cours d'élaboration pour vous aider à les localiser et à les évaluer (voir les sections Outils et Ressources)

Outils de migration

Généralement, il s'agit d'outils qui transforment un format de fichier obsolète en un format plus récent qui peut être pris en charge. De nombreux outils permettent d'effectuer cette migration. Ils ont tendance à se limiter à un seul aspect (par exemple, ImageMagick ne fonctionne que pour les fichiers d'images numériques).
Un outil de migration n'est qu'une partie d'une trajectoire de migration. La trajectoire doit comprendre un format de destination/cible, que vous aurez sélectionné conformément aux indications données ci-dessus.

Les outils de migration peuvent présenter des risques. L'un de ces risques est la modification « invisible » du contenu ou des données lors de la migration. Pour réduire ce risque, une stratégie consiste à définir un ensemble de critères d'acceptation des éléments que l'objet transformé doit conserver, par exemple en termes de formatage, de présentation, ou même de fonctionnalité, et à confirmer les résultats souhaités par un processus d'assurance qualité.

La migration des formats de fichiers n'est pas toujours la solution. Certains formats de fichiers CAO (conception assistée par ordinateur, NdT) et FAO (fabrication assistée par ordinateur, NdT), par exemple, ne peuvent pas être facilement transférés. L'industrie aérospatiale a constaté que la migration des anciens fichiers de CAO vers un format plus récent nécessite beaucoup de validations, principalement parce qu'elle est tenue par un cadre réglementaire de démontrer que leurs données restent cohérentes et répondent à des normes très strictes. En bref, le coût de la migration et de la validation est (pour ce secteur) beaucoup plus élevé qu'une solution d'émulation, une approche qui (dans ce cas) implique de conserver le logiciel de CAO et de le maintenir.
Voir également les sections Outils et Préservation du contenu spécifique.

Outils de restitution

Généralement, il s'agit d'outils qui peuvent lire et rejouer un format de fichier, afin que la communauté des utilisateurs puisse lire et interpréter la ressource ; cela s'applique le plus souvent aux fichiers stockés dans des formats accessibles. Un exemple d'outil de rendu de base est PDF Reader. Un outil de rendu plus sophistiqué pourrait être le lecteur multimédia de la *Wellcome Library*, qui prend en charge les textes issus de l'OCR, les images et les fichiers audiovisuels.

Outils d'identification du format des fichiers

Certains outils peuvent identifier des caractéristiques des formats de fichiers qui ne sont pas immédiatement évidents à partir de leur extension de fichier en lisant l'en-tête du format de fichier. Ils peuvent ainsi identifier par exemple le type, la taille, la version. Parmi les exemples de tels outils, on peut citer PRONOM, JHOVE et l'outil d'extraction de métadonnées de la Nouvelle-Zélande (voir les ressources ci-dessous).

Ces outils sont utilement déployés au moment de l'ingestion afin que vous sachiez dès le départ quel type de format de fichier vous importez dans l'archive.

Certains outils d'identification peuvent également indiquer des outils de restitution susceptibles d'être pertinents, ou même (comme PRONOM) suggérer une trajectoire de migration basée sur l'identification du format de fichier.

Outils de validation du format des fichiers

JHOVE est l'un des rares outils capables de valider un format de fichier. Il le fait en comparant une instance d'un format de fichier avec des ensembles de profils attendus, enregistrés dans sa librairie informatique. JHOVE peut créer des rapports sur certains formats de fichiers et spécifier s'ils sont valides et bien structurés.

Étude de caractérisation d'une collection

Il s'agit d'une étude sur les formats de fichiers utilisés afin de savoir ce dont vous disposez et de caractériser vos collections. Celle-ci est liée à une stratégie de planification, vous permettant de savoir ce que vous devez maintenir, et l'effort nécessaire pour le faire.

Une telle étude doit accorder une attention particulière aux versions des formats de fichiers, et aux logiciels nécessaires à leur lecture et restitution. Si possible, rassemblez des informations sur les spécifications publiées pour ces formats ; certaines spécifications sont publiées sur le web.

Des travaux utiles ont été réalisés dans ce domaine à la *British Library*, avec des projets sur l'évaluation de la durabilité (Maureen Pennock, Paul Wheatley, Peter May) et le profilage des collections (Michael Day, Maureen Pennock, Ann MacDonald). Au moment de la rédaction du présent document, il n'existe pas de liens actifs vers ces projets, mais il est prévu que les travaux sur l'évaluation de la durabilité soient publiés sur le Wiki de la DPC. Ces approches sont utiles et peuvent être considérées comme des exemples des meilleures pratiques actuelles. Même si vous ne procédez pas à une évaluation ou à un profilage aussi approfondi que celui de la BL, l'exercice est utile et réalisable en pratique.

Éviter la prolifération des types de fichiers

Dans la mesure du possible, limitez l'éventail des formats de fichiers que vous prenez en charge, afin d'en réduire la complexité. Une bonne approche de la planification de la préservation consiste à normaliser, plutôt qu'à ajouter de multiples formats de migration à votre collection. Plus la gamme de formats est réduite, plus la charge de prise en compte sera faible.

Communauté

Il s'agit d'identifier un consensus autour des formats de fichiers cibles et de collaborer avec les institutions qui détiennent des collections similaires à la vôtre. Avec quels formats choisissent-elles de travailler ?

Conclusion

Pour certains types de contenus, il existe un consensus sur le choix du format de préservation. Par exemple, l'archivage audio où le format WAV est couramment utilisé. Dans d'autres domaines, le consensus est beaucoup plus difficile à atteindre. La préservation de la vidéo numérique est un domaine complexe où les progrès ont été entravés par un manque de consensus et une prolifération incontrôlée de formats d'empaquetage, de méthodes de livraison et de méthodes d'encodage. Le choix des formats de fichiers d'images est légèrement plus simple, avec un choix limité de formats pour l'archivage et d'autres pour la livraison. Il a été généralement admis que le format TIFF est le format correct pour l'archivage des fichiers master (le format RAW ou DNG peut aussi être utilisé pour l'archivage) mais cela est maintenant remis en question par le format JPEG2000 qui offre un niveau de compression sans perte bien plus élevé que TIFF et qui bénéficie d'un code Open Source.

Ressources

Library of Congress recommended format specifications

http://www.loc.gov/preservation/resources/rfs/index.html

La Bibliothèque du Congrès a élaboré un ensemble de spécifications de formats qu'elle recommande, tant en interne pour ses propres équipes qu'en externe pour les créateurs, les vendeurs et les archivistes, comme étant les formats à utiliser de préférence pour assurer la préservation et l'accès à long terme. Elle couvre les formats numériques et analogiques et se divise en six grandes catégories : œuvres textuelles et compositions musicales ; œuvres d'images fixes ; œuvres sonores ; œuvres d'images animées ; logiciels et jeux ludiques ou éducatifs ; et ensembles de données/bases de données.

Rapports du Jisc sur les propriétés significatives

Entre 2007 et 2008, le Jisc a financé cinq études de propriétés significatives pour différents types de contenus et de fichiers. Notez que la discussion dans les rapports est datée de 2007-2008. Les rapports sont les suivants :

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

L'équipe de préservation numérique de la *British Library* a entrepris des évaluations des formats de fichiers présentant un risque de préservation afin d'acquérir des connaissances sur les lacunes des meilleures pratiques actuelles, la compréhension et la capacité à travailler avec des formats de fichiers spécifiques. Chaque évaluation se concentre sur la saisie des risques de préservation fondés sur des preuves et sur les implications de l'obsolescence institutionnelle qui entraîne des problèmes de maintenance du contenu au fil du temps. Les évaluations sont hébergées dans une nouvelle section sur le Wiki de la DPC. Trois évaluations couvrant les formats JP2, TIFF et PDF ont débuté la série.

Library of Congress Sustainability Factors

http://www.digitalpreservation.gov/formats/index.shtml

Ce site s'intéresse aux formats associés aux contenus numériques indépendants des supports, c'est-à-dire les contenus qui sont généralement gérés sous forme de fichiers et ne dépendent pas d'un support physique particulier. Il ne concerne pas les formats associés au contenu numérique dépendant des médias, c'est-à-dire les formats qui dépendent et sont inextricablement liés à un support physique, par exemple les DVD, les CD audio et les formats de bandes vidéo comme DigiBeta. Il identifie et décrit les formats qui sont prometteurs pour la durabilité à long terme, et développe des stratégies pour maintenir ces formats, y compris des recommandations concernant les outils et la documentation nécessaires à leur gestion.

Help Solve the File Format Problem

http://fileformats.archiveteam.org

Un Wiki d'information sur les formats de fichiers, conçu à l'aide du crowdsourcing, sur le site de l'équipe des archives. Tout le contenu est disponible sous une licence Creative Commons 0.

Is JPEG 2000 a digital preservation risk?

http://blogs.loc.gov/digitalpreservation/2013/01/is-jpeg-2000-a-preservation-risk/

Un blog intéressant et un fil de discussion sur le format d'image JPEG 2000.

OPF File Format Risk Registry

http://wiki.opf-labs.org/display/TR/OPF+File+Format+Risk+Registry

Il s'agit plus particulièrement des questions et des risques liés aux formats de fichiers qui ont des implications pour la préservation et l'accessibilité à long terme et de la manière de les traiter de façon pratique. Il vise à compléter les registres de formats plus formels.

PRONOM

http://apps.nationalarchives.gov.uk/pronom/Default.aspx

Ce registre de formats de fichiers est une ressource majeure pour toute personne ayant besoin d'informations impartiales et définitives sur les formats de fichiers, les logiciels et autres composants techniques nécessaires pour soutenir l'accès à long terme aux documents électroniques et autres objets numériques ayant une valeur culturelle, historique ou commerciale.

DROID (Digital Record Object Identification)

http://www.nationalarchives.gov.uk/information-management/manage-information/preserving-digital-records/droid/

Il s'agit d'un outil d'identification automatique des formats de fichiers qui fournit des catégories d'identification de format pour les fichiers inconnus dans une collection numérique. Il utilise des signatures internes pour identifier et signaler le format de fichier et la version spécifiques des fichiers numériques. Ces signatures sont stockées dans un fichier de signature XML, généré à partir des informations enregistrées dans le registre PRONOM.

Études de cas

Voir la section Études de cas détaillées sur la préservation du contenu du manuel pour les études de cas pertinentes

Sécurité de l'information

Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Introduction

Cette section est destinée à guider les professionnels de niveau débutant ou intermédiaire sur les implications de la sécurité de l'information pour la préservation numérique. Les questions de sécurité de l'information concernent la sécurité du système (par exemple, la protection de la préservation numérique et des systèmes/services en réseau contre l'exposition à des menaces externes/internes) ; la sécurité des collections (par exemple, la protection des données contre la perte ou la modification, l'autorisation et l'audit des processus de dépôt) ; et les aspects juridiques et réglementaires (par exemple, les informations personnelles ou confidentielles contenues dans les objets numériques, l'accès sécurisé, la rédaction). La sécurité de l'information est un sujet complexe et important pour les systèmes d'information en général. Il est important de s'appuyer sur une expertise pertinente au sein de votre organisation et au-delà, par le biais du gouvernement et d'autres réseaux, pour les procédures et conseils généraux en matière de sécurité de l'information. Vous pouvez également avoir besoin d'une déclinaison appropriée pour des procédures et des exigences spécifiques de préservation numérique.

Des procédures de sécurité rigoureuses permettront :

1. D'assurer le respect de toutes les exigences légales et réglementaires ;
2. De protéger les objets numériques contre les modifications accidentelles ou délibérées ;
3. De fournir un journal d'audit pour satisfaire aux exigences en matière de responsabilité ;
4. D'agir comme un moyen de dissuasion contre les éventuelles atteintes à la sécurité interne ;
5. De protéger l'authenticité des documents numériques ;
6. De se protéger contre le vol ou la perte.

De nombreux types d'objets numériques sélectionnés pour une préservation à long terme peuvent contenir des informations confidentielles et sensibles qui doivent être protégées pour garantir qu'elles ne sont pas accessibles à des utilisateurs non autorisés. Dans de nombreux cas, il peut s'agir d'obligations légales ou réglementaires imposées à l'organisation. Ces objets doivent être gérés conformément à la politique de sécurité de l'information de l'organisation afin de les protéger contre les atteintes à la sécurité. La norme ISO 27001 décrit la manière dont les procédures de sécurité peuvent être codifiées et contrôlées (ISO, 2013a). La norme ISO 27002 fournit des lignes directrices sur la mise en œuvre de procédures de sécurité conformes à la norme ISO 27001 (ISO, 2013b). Les organisations conformes peuvent être accréditées et validées par des organismes externes. Dans certains cas, la politique de sécurité de l'information de votre propre organisation peut également avoir un impact sur les activités de préservation numérique et vous

devrez peut-être faire appel au soutien de vos équipes de gouvernance de l'information et des technologies de l'information et de la communication pour faciliter vos processus.

Les méthodes de sécurité de l'information telles que le cryptage ajoutent à la complexité du processus de préservation et doivent être évitées autant que possible lors des réplications. D'autres méthodes de sécurité plus rigoureuses pourraient donc être nécessaires pour les fichiers sensibles non chiffrés, par exemple restreindre l'accès aux terminaux verrouillés dans des lieux contrôlés (salles sécurisées), ou exiger une authentification forte des utilisateurs pour l'accès à distance. Toutefois, ces autres approches ne sont pas toujours suffisantes ou réalisables. Le cryptage peut également être présent sur les fichiers qui sont reçus par versement d'un déposant, il est donc important d'être conscient des options de sécurité de l'information telles que le cryptage, la gestion des clés de cryptage, et leurs implications pour la préservation numérique.

Techniques de protection de l'information

Plusieurs techniques de sécurité de l'information peuvent être appliquées pour protéger l'objet numérique :

Cryptage

Le cryptage est une technique cryptographique qui protège l'objet numérique en le convertissant en une forme codée. Le cryptage peut être appliqué à plusieurs niveaux, d'un simple fichier à un disque entier. Il existe de nombreux algorithmes de cryptage, chacun d'entre eux codant les informations d'une manière différente. Ils nécessitent l'utilisation d'une clé pour décrypter les données et les reconvertir dans leur forme originale. La puissance de la méthode de cryptage est liée à la taille de la clé. Par exemple, un cryptage de 256 bits sera plus sûr qu'un cryptage de 128 bits.

Il convient de noter que le cryptage n'est efficace que lorsqu'un tiers n'a pas accès à la clé de cryptage utilisée. Un utilisateur qui a entré le mot de passe d'un lecteur crypté et qui a laissé sa machine allumée et sans surveillance donnera à des tiers la possibilité d'accéder aux données détenues dans la zone cryptée, ce qui peut entraîner leur fuite.

De même, les mesures de sécurité du cryptage (si elles sont utilisées) peuvent perdre de leur efficacité avec le temps dans un dépôt : il y a effectivement une course aux armements entre les techniques de cryptage et les méthodes de calcul pour les briser. Par conséquent, s'il est utilisé, tout cryptage par un référentiel doit être activement géré et mis à jour au fil du temps pour rester sûr.

L'objet numérique crypté ne peut être accessible dans le temps dans un dépôt d'archives que si l'organisation gère ses clés. La perte ou la destruction de ces clés rendra les données inaccessibles.

Contrôles d'accès

Les contrôles d'accès permettent à un administrateur de spécifier qui est autorisé à accéder à l'objet numérique et le type d'accès qui est autorisé (par exemple, lecture seule, écriture). Le manuel suit les niveaux de préservation de la *National Digital Stewardship Alliance* (NDSA) en recommandant quatre niveaux auxquels la préservation numérique peut être soumise pour le contrôle d'accès. Les niveaux de la NDSA visent principalement à comprendre qui a accès au contenu, qui peut effectuer quelles actions sur ce contenu et à faire respecter ces restrictions d'accès (NDSA, 2013) comme suit :

Niveau NDSA	Activité

1	• Identifier qui a l'autorisation pour lire, écrire, déplacer ou détruire des fichiers particuliers • Restreindre les personnes ayant ces autorisations sur des fichiers particuliers
2	• Documenter les restrictions d'accès au contenu
3	• Maintenir des journaux sur les personnes qui réalisent des actions sur les fichiers, y compris les suppressions et les actions de préservation
4	• Réaliser des audits des journaux

Suppression

La suppression fait référence au processus d'analyse d'une ressource numérique, d'identification d'informations confidentielles ou sensibles, et de suppression ou de remplacement de ces informations. Les techniques couramment utilisées comprennent l'anonymisation et la pseudonymisation pour supprimer les informations permettant d'identifier une personne, ainsi que le nettoyage des informations relatives à l'auteur. Lorsqu'il s'agit d'ensembles de données, cette opération est généralement effectuée en supprimant des informations tout en conservant la structure de l'enregistrement dans la version diffusée. Vous devez toujours procéder aux modifications sur une copie de l'original, jamais sur l'original lui-même.

La majorité des documents numériques créés à l'aide de systèmes de bureautique, tels que Microsoft Office, sont stockés dans des formats propriétaires à codage binaire. Les formats binaires peuvent contenir des informations importantes qui ne sont pas affichées, et leur présence peut donc ne pas être apparente. Ils peuvent incorporer des historiques de modifications, des suivis d'audit ou des métadonnées intégrées, grâce auxquelles les informations supprimées peuvent être récupérées ou de simples processus d'édition facilement contournés. Les documents numériques peuvent être expurgés par une combinaison de suppression d'informations et de conversion dans un format différent. Certains formats, tels que les fichiers texte ASCII, ne contiennent que des informations affichables. La conversion à ce format éliminera donc toute information qui pourrait être cachée dans des parties non affichables d'un flux binaire.

Ressources

ENISA. 2013, Cloud Security Incident Reporting
https://www.enisa.europa.eu/activities/Resilience-and-CIIP/cloud-computing/incident-reporting-for-cloud-computing/
L'Agence européenne chargée de la sécurité des réseaux et de l'information propose des recommandations sur la manière dont les fournisseurs de services dans le cloud et leurs clients doivent réagir aux violations de la sécurité et les signaler. (38 pages).

ISO 27001:2013 - Technologies de l'information- Techniques de sécurité - Systèmes de gestion de la sécurité de l'information - Exigences. Genève : Organisation internationale de normalisation
http://www.iso.org/iso/catalogue_detail?csnumber=54534

La norme ISO 27001 décrit la manière dont les procédures de sécurité peuvent être codifiées et contrôlées. Les organisations conformes peuvent être accréditées et validées à l'extérieur. Un modèle pour un ensemble de politiques alignées sur la norme est disponible. Notez qu'il s'agit de rubriques, destinées à faciliter la création de politiques, plutôt que de déclarations de politiques. Toutefois, des ensembles de politiques similaires sont utilisés dans un grand nombre d'organisations. (23 pages).

Version française : https://www.iso.org/fr/standard/54534.html

ISO 27002:2013, Technologies de l'information - Techniques de sécurité - Code de pratique pour les contrôles de sécurité de l'information. Genève : Organisation internationale de normalisation
http://www.iso.org/iso/catalogue_detail?csnumber=54533
La norme ISO 27002 fournit des lignes directrices sur la mise en œuvre de procédures de sécurité conformes à la norme ISO 27001. (80 pages)

Version française : https://www.iso.org/fr/standard/54533.html

ISO 27799:2008 - *Informatique de santé - Gestion de la sécurité de l'information dans le domaine de la santé* à l'aide d'ISO/IEC 27002. Genève : Organisation internationale de normalisation
http://www.iso.org/iso/catalogue_detail?csnumber=41298
La norme ISO 27799 fournit des conseils spécifiques sur la mise en œuvre des normes ISO 27002 et 27001 dans le secteur de la santé. (58 pages)

Norme révisée par l'ISO 27799:2016 *Informatique de santé - Management de la sécurité de l'information relative à la santé en utilisant l'ISO/IEC 27002*. Version française disponible à l'adresse suivante : https://www.iso.org/fr/standard/62777.html

Cabinet Office, 2009, HMG IA Standard No. 1 - Technical Risk Assessment
https://www.ncsc.gov.uk/guidance/information-risk-management-hmg-ia-standard-numbers-1-2
Une discussion détaillée et une norme destinées aux gestionnaires de risques et aux professionnels de l'assurance de l'information du Royaume-Uni qui sont chargés d'identifier, d'évaluer et de traiter les risques techniques pour les systèmes et les services qui traitent, stockent et traitent les informations numériques du gouvernement. (114 pages).

Boîte à outils pour la suppression (Archives nationales du Royaume-Uni, 2011)
http://www.nationalarchives.gov.uk/documents/information-management/redaction_toolkit.pdf
Cette boîte à outils des Archives nationales du Royaume-Uni a été produite en 2011 pour fournir des conseils sur l'édition d'objet exempté de l'information détenue par les organismes publics. Elle couvre les principes génériques des documents sur tous les supports, mais comporte une petite section spécifiquement consacrée aux documents électroniques et des conseils détaillés sur les méthodes d'édition sécurisée des documents électroniques de tous types(21 pages).

BitCurator
https://bitcurator.net/

BitCurator est une suite d'outils Open Source d'analyse de données et de criminalistique numérique pour aider les institutions de collecte d'objets numériques. Certains de ces outils permettent de localiser des informations privées et sensibles sur des supports numériques et de préparer des documents pour un accès public.

Bureau des commissaires à l'information (ICO) : Sécurité de l'information
https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/security/
Le site web de l'ICO contient des conseils sur le signalement des infractions à la sécurité et l'utilisation des technologies de l'information. Pour les personnes travaillant dans des organisations relevant de la compétence de l'ICO, il est essentiel de comprendre les recommandations de ces lignes directrices afin d'entamer des conversations avec les collègues des TIC et de la gouvernance de l'information, car ils devront être assurés que les travaux peuvent être effectués conformément aux recommandations de l'ICO.

"Access to the Secure Lab"
https://www.ukdataservice.ac.uk/get-data/how-to-access/accesssecurelab
Un certain nombre de sources de microdonnées confidentielles et sensibles deviennent disponibles par le biais de datalabs à travers le Royaume-Uni. Ces données sont considérées comme potentiellement identifiables et ne peuvent être consultées que par le biais d'un laboratoire de données (par opposition au téléchargement). En outre, les chercheurs sont invités à remplir un certain nombre d'exigences supplémentaires. Certaines de ces données sont accessibles via le laboratoire sécurisé du service de données du Royaume-Uni et cette page fournit des aperçus utiles et un accès aux accords d'utilisation pertinents.

Études de cas

Ouverture de l'accès aux données administratives pour l'évaluation des services publics : Le cas du Laboratoire de données sur la justice
http://evi.sagepub.com/content/21/2/232.full.pdf+html
Le Laboratoire de données sur la justice, une unité située dans un cadre sécurisé et détenant une expertise en matière d'évaluation et de statistiques, a permis aux fournisseurs de programmes visant à réduire la récidive d'obtenir des preuves sur la manière dont l'impact de leurs interventions diffère de celui d'un groupe de comparaison apparié. Cet article explore le développement du *Justice Data Lab*, les défis méthodologiques et autres rencontrés, et les expériences des organisations d'utilisateurs. L'article tire les conséquences pour le futur développement des Data Labs et l'utilisation des données administratives pour l'évaluation des services publics (16 pages).

UK Data Service: Data Security
https://www.ukdataservice.ac.uk/manage-data/store/security

Cette page web résume la manière dont *UK Data Archive* gère la sécurité des données de leurs fonds. La sécurité des données peut être nécessaire pour protéger les droits de propriété intellectuelle, les intérêts commerciaux ou pour assurer la sécurité des informations sensibles. Les dispositions prises doivent être proportionnées à la nature des données et aux risques encourus. Une attention particulière à la sécurité est également nécessaire lorsque les données doivent être détruites.

Références

NDSA, 2013. *The NDSA Levels of Digital Preservation: An Explanation and Uses, version 1* (2013). Disponible à l'adresse suivante : http://www.digitalpreservation.gov/ndsa/working_groups/documents/NDSA_Levels_Archiving_2013.pdf

Traduction française de la v2 : Thomas Ledoux, Marion Humbert, Edouard Vasseur. Utilisation des Niveaux de Préservation Numérique : aperçu de la V2.0, traduction française. 2020 https://hal-bnf.archives-ouvertes.fr/hal-02551807/document

ISO, 2013a. *ISO 27001:2013 - Technologies de l'information – Techniques de sécurité - Systèmes de management de la sécurité de l'information - Exigences*. Geneva : *International Organization for Standardization*. Disponible à l'adresse suivante : http://www.iso.org/iso/catalogue_detail?csnumber=54534

Version française : https://www.iso.org/fr/standard/54534.html

ISO, 2013b. *ISO 27002:2013 – Technologies de l'information – Techniques de sécurité – Code de bonne pratique pour le management de la sécurité de l'information*. Geneva : *International Organization for Standardization*. Disponible à l'adresse suivante : http://www.iso.org/iso/catalogue_detail?csnumber=54533

Version française : https://www.iso.org/fr/standard/54533.html

Services *cloud*

Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Qu'est-ce que le *cloud computing* ?

Le *cloud computing* est un terme qui englobe un large éventail de cas d'utilisation et de modèles de mise en œuvre. En substance, un *cloud* informatique est un vaste ensemble de ressources informatiques partagées, y compris le stockage de données. Lorsqu'une personne a besoin d'une puissance de calcul supplémentaire, elle peut simplement l'extraire de l'ensemble sans grand effort manuel (et souvent sans aucun effort) de la part de l'équipe informatique, ce qui réduit les coûts et raccourcit considérablement le temps nécessaire pour commencer à utiliser de nouvelles ressources informatiques. La plupart de ces *clouds* sont gérés sur l'Internet public par des sociétés bien connues comme Amazon et Google. Certaines grandes organisations ont également trouvé un intérêt à exploiter des *clouds* privés dans leurs propres centres informatiques, où des économies d'échelle similaires commencent à s'appliquer.

Les caractéristiques généralement acceptées d'un service de *cloud computing* typique peuvent être définies comme étant les ordinateurs et le stockage de données qui sont :

- Disponibles au moment voulu (« à la demande »), sans qu'il soit nécessaire de recourir à de longs processus d'achat et de configuration ;
- Disponibles sur des réseaux standard tels qu'Internet, sans exigences particulières pour les réseaux, protocoles ou matériels obscurs ou propriétaires ;
- Capable d'offrir une capacité supplémentaire lorsque la demande augmente, et moindre lorsque la demande diminue (« élasticité ») ;
- Capable de ne facturer aux clients que le stockage qu'ils utilisent.

Cloud computing et préservation numérique

Le *cloud computing* peut offrir plusieurs avantages :

- La flexibilité du *cloud* permet de tester et d'expérimenter de manière relativement rapide et peu coûteuse les nouveaux fournisseurs de services. Il existe déjà quelques activités pilotes avec ces services *cloud* et des possibilités d'apprentissage partagé au sein de la communauté ;

- Il y a maintenant une plus grande flexibilité et plus d'options dans le déploiement des services de stockage dans le *cloud* et donc une plus grande pertinence pour les dépôts d'archives par rapport aux années précédentes (voir *Cloud* public, communautaire, privé et hybride) ;
- Des économies de coûts sont possibles grâce à la simplification des achats et aux économies d'échelle, en particulier pour les petits dépôts d'archives. Ces éléments sont importants en cette période de pressions financières ;
- Les services *cloud* peuvent fournir une réplication facile et automatisée vers plusieurs sites, essentielle pour la planification de la reprise d'activité et l'accès à un stockage numérique géré par des professionnels ; en outre, les spécialistes peuvent ajouter l'accès à d'autres outils, procédures, processus de travail et accords de service dédiés, adaptés aux exigences de la préservation numérique.

Modèles de services *cloud* et fournisseurs de services

Il existe quatre modèles différents de services *cloud* :

- Public - Services commerciaux hébergés dans de grands centres de calcul dans le monde entier, accessibles sur les réseaux publics à toute personne ayant les moyens de payer.
- Privé - Les grandes organisations créent leur propre *cloud* en virtualisant de grands ensembles de serveurs physiques dans leurs propres centres de calcul.
- Hybride - Combinaison de plusieurs aspects du *cloud* public et privé, généralement pour gérer de grandes fluctuations dans la demande ou pour répondre à différentes exigences de sécurité.
- Communautaire - Sur le plan architectural, il peut être identique à un service de *cloud* public, mais optimisé pour un groupe particulier d'utilisateurs dont l'accès est limité.

Il existe actuellement deux catégories de fournisseurs de services *cloud* : les généralistes qui offrent un stockage dans le *cloud* (Amazon, Rackspace, Google, etc.) et les entreprises spécialisées qui répondent à des exigences et des fonctions de préservation numérique spécifiques supplémentaires (voir Ressources et études de cas pour des exemples).

Avantages

- Les services *cloud* peuvent fournir une réplication facile et automatisée vers plusieurs endroits et un accès à un stockage numérique géré par des professionnels et dont l'intégrité est contrôlée. En conséquence, la préservation des bits (pérennité) de l'information numérique peut être au moins aussi bonne (ou meilleure) que celle qui peut être obtenue localement ;
- Les services d'archives peuvent ajouter l'accès à des outils, des procédures, des processus de travail et des accords de service dédiés, adaptés aux exigences de la préservation numérique par l'intermédiaire de fournisseurs spécialisés ;
- La simplification des procédures d'achat et les économies d'échelle, notamment pour les petits services d'archives, peuvent permettre de réaliser des économies ;
- La flexibilité du *cloud* permet de tester et d'expérimenter les fournisseurs de manière relativement rapide et peu coûteuse ;
- Il existe une plus grande flexibilité et davantage d'options dans le déploiement des services dans le *cloud* et donc une plus grande pertinence pour les archives par rapport aux années précédentes. En particulier, les mises en œuvre de *cloud* privé ou de *cloud* hybride peuvent répondre aux préoccupations de sécurité concernant le stockage d'objet plus sensibles qui peut être considéré comme inapproprié pour le *cloud* public ;

- Des stratégies de sortie peuvent être mises en place pour répondre aux préoccupations des services d'archives concernant la stabilité et la longévité des fournisseurs ou d'autres risques liés au changement. Par exemple, la synchronisation du contenu entre deux fournisseurs de services dans le *cloud* ou entre un *cloud* externe et un stockage interne local ; ou encore la mise en œuvre d'une copie sous séquestre détenue indépendamment par un tiers de confiance ;
- Il existe déjà quelques activités en expérimentation avec ces services *cloud* et des possibilités d'apprentissage partagé au sein de la communauté.

Inconvénients

- Le *cloud* est conçu pour la flexibilité et le changement rapide. Les services d'archives ont plutôt une vision à long terme. Le stockage dans le *cloud* et les contrats de service doivent être gérés avec soin au fil du temps pour répondre aux besoins des archives. Il faut s'attendre à ce que les données conservées dans les dépôts d'archives soient à la fois préservées et accessibles au-delà de la durée de vie commerciale de toute technologie ou de tout fournisseur de services actuel ;
- Le *cloud* peut être moins cher, mais il oblige souvent les organisations à penser différemment la gestion de leurs budgets. Il mobilise également des compétences différentes de celles qu'on a coutume de trouver chez les services d'achat et gestionnaires de contrats de services informatiques, ce qui peut impliquer des coûts de formation ou de recrutement ;
- Les services de *cloud* public ont tendance à facturer chaque mois la capacité qui a été réellement consommée. Il peut donc être difficile d'établir un budget à l'avance ou de prévoir avec précision la quantité de données susceptibles d'être téléchargées, stockées ou téléchargées (cependant, certains fournisseurs peuvent vous facturer un abonnement annuel basé sur le volume) ;
- Comme pour toute forme d'externalisation, il est important que les services d'archives fassent preuve de discernement dans l'évaluation et le contrôle des risques du stockage dans le *cloud*. Assurez-vous que toutes les exigences et obligations légales relatives aux droits des tiers sur les données à stocker seront respectées. Celles-ci peuvent concerner la gestion, la préservation ou l'accès, et peuvent avoir été imposées aux services d'archives et à leurs organisations mères par leurs donateurs et financeurs par le biais de contrats et d'accords ou par la législation du gouvernement ;
- L'utilisation des services *cloud* exigera des services d'archives qu'ils examinent les questions liées aux droits d'auteur, notamment : qui détient actuellement les droits d'auteur ; si des autorisations de licence supplémentaires peuvent être nécessaires ; quelles autorisations le fournisseur de services *cloud* devra-t-il obtenir pour fournir le service ; si le fournisseur de services cloud est en mesure d'utiliser les données à ses propres fins ; et quelle partie détiendra les droits sur les données ou les œuvres créées à partir des données originales ;
- L'utilisation des services *cloud* peut soulever des questions de sécurité des données, lorsque les données concernées sont des « données à caractère personnel » (par exemple, des données qui permettent l'identification d'une personne vivante), notamment en ce qui concerne la détermination de la responsabilité de la sécurisation des données et l'audit des fournisseurs, ainsi que le lieu de traitement et la mesure dans laquelle les risques encourus par l'automatisation de la fourniture de services peuvent être traités par contrat ;
- Les éléments juridiques de la relation entre un entrepôt numérique et un ou plusieurs fournisseurs de services *cloud* (par exemple, les conditions des contrats de service et les accords de niveau de service) doivent être bien définis et répondre à vos exigences. Cela peut être difficile car de nombreux fournisseurs de services *cloud* ont des accords de niveau

de service et des contrats standard pour obtenir des prix de base et ont une flexibilité limitée pour négocier les conditions ;
- Des dispositions explicites doivent être prises pour des stratégies de sortie prédéfinies et des procédures de test, de contrôle et d'audit efficaces.

Conclusions

Le terme *cloud* peut englober un large éventail de modèles de mise en œuvre des services de préservation numérique. Il y a beaucoup à apprendre des organisations qui ont déjà expérimenté ou adopté l'utilisation du *cloud*. Par exemple, plusieurs services d'archives ont pu répondre aux préoccupations les plus répandues concernant les services *cloud* et trouver des moyens d'intégrer avec succès le stockage *cloud* dans leurs activités de préservation numérique. D'autres utilisent les services *cloud* pour tout ou partie de leurs autres fonctions de préservation numérique, telles que la planification de la préservation. En somme, l'acquisition de services *cloud* est similaire à l'acquisition de toute technologie de l'information. Vous devez gérer et traiter les risques comme vous le feriez pour n'importe quelle autre partie de votre infrastructure informatique.

Ressources

Guide des Archives nationales du Royaume-Uni sur le stockage *cloud* et la préservation numérique (2e édition 2015)
http://www.nationalarchives.gov.uk/documents/CloudStorage-Guidance_March-2015.pdf
Ce guide explore l'évolution du stockage *cloud* dans la préservation numérique, les nouvelles options et les bonnes pratiques, ainsi que les exigences et les normes que les archives devraient prendre en compte. Des sections sont consacrées aux services, aux questions juridiques et à cinq études de cas liées entre elles. Des sources de conseils et d'orientations supplémentaires sont également incluses (39 pages).

Aitken, B, McCann, P, McHugh, A and Miller, K, 2012, Digital Curation and the Cloud, DCC
http://www.dcc.ac.uk/sites/default/files/documents/Curation-in-the-Cloud_master_final.pdf
Ce rapport de 2012 se concentre sur l'utilisation des services *cloud* pour la préservation des données de recherche. Il fournit quelques définitions du *cloud computing* et examine un certain nombre d'approches du *cloud* ouvertes aux établissements d'enseignement supérieur en 2012 (30 pages).

Anderson. S, 2014, Feet On The Ground : Une approche pratique du *cloud*. Neuf éléments à prendre en compte lors de l'évaluation du stockage *cloud*, de la préservation de l'audiovisuel
http://www.avpreserve.com/wp-content/uploads/2014/02/AssessingCloudStorage.pdf
Un livre blanc sur les services *cloud*, divisé en neuf sujets et questions à poser. Des profils de fournisseurs par rapport à ces neuf sujets sont disponibles (7 pages).

A. Brown, C. Fryer, "*Achieving Sustainable Digital Preservation in the Cloud*" (Parvenir à une préservation numérique durable dans le *cloud*)
http://www.girona.cat/web/ica2014/ponents/textos/id87.pdf
Ce document décrit comment le Parlement utilise le *cloud* dans le cadre de son infrastructure de dépôt numérique. 2004 (10 pages).

Spécialiste de la préservation numérique Fournisseurs de services *cloud*

ArchivesDirect
http://archivesdirect.org
ArchivesDirect propose une instance hébergée d'Archivematica avec un stockage via DuraCloud en Amazon S3 et Amazon Glacier sécurisé et répliqué.

Arkivum
http://arkivum.com
L'*Archive as a Service* d'Arkivum fournit un service entièrement géré et sécurisé pour la préservation des données à long terme avec un accès en ligne et une garantie d'intégrité des données qui fait partie de son accord de niveau de service et qui est soutenu par une assurance mondiale.

DuraCloud
http://www.duracloud.org
DuraCloud est un service géré de *DuraSpace*. Il fournit une assistance et des outils qui copient automatiquement le contenu sur plusieurs fournisseurs de stockage *cloud* différents et garantit que toutes les copies du contenu restent synchronisées. Voir également ArchivesDirect pour son service commun avec Archivematica.

Preservica
https://preservica.com/digital-archive-software/products-editions
Preservica Cloud Edition est une plateforme de préservation numérique entièrement hébergée dans le *cloud*, conforme à OAIS, qui comprend également un accès/découverte public pour vous permettre de partager vos archives ou votre collection en toute sécurité.

Le blog de David Rosenthal
http://blog.dshr.org/
Contient un certain nombre d'articles sur l'économie du *cloud computing*.

Études de cas

Étude de cas des Archives nationales du Royaume-Uni : Groupe de travail sur la préservation numérique du Conseil des archives et des dossiers du Pays de Galles
http://www.nationalarchives.gov.uk/documents/Cloud-Storage-casestudy_Wales_2015.pdf
Cette étude de cas examine l'expérience d'un groupe de travail intersectoriel d'archives galloises qui a coopéré pour tester une série de systèmes et de déploiements de services dans le cadre d'une validation de principe pour l'archivage dans le *cloud*. Elle explique le contexte organisationnel, la nature variée de leurs exigences et approches en matière de préservation numérique, et leur expérience en matière de sélection, de déploiement et de test de la préservation numérique dans le cloud. L'étude de cas a examiné le logiciel open source Archivematica avec Windows Azure de Microsoft ; Archivematica avec CloudSigma ; *Preservica Cloud Edition* et a commencé à tester Archivematica avec Arkivum 100. Janvier 2015 (10 pages).

Étude de cas des Archives nationales du Royaume-Uni : *Tate Gallery*
http://www.nationalarchives.gov.uk/documents/Cloud-Storage-casestudy_Tate_Gallery_2015.pdf
Cette étude de cas traite de l'expérience de développement d'un entrepôt numérique partagée pour les quatre sites physiques de la Tate, alimentée par un système de stockage commercial d'Arkivum. Elle explique le contexte organisationnel, la nature de leurs besoins et approches en matière de préservation numérique, et les raisons pour lesquelles ils ont choisi la solution sur site d'Arkivum, Arkivum/OnSite, de préférence à toute offre basée sur le *cloud*. Il se termine par les principaux enseignements tirés et examine les plans des futurs développements. Janvier 2015 (7 pages).

Étude de cas des Archives nationales du Royaume-Uni : *Dorset History Centre*
http://www.nationalarchives.gov.uk/documents/Cloud-Storage-case-study_Dorset_2015_%281%29.pdf
Cette étude de cas porte sur le *Dorset History Centre*, un service d'archives du gouvernement local. Elle explique le contexte organisationnel de l'Archive, la nature de ses exigences et approches en matière de préservation numérique, son projet pilote de deux ans utilisant *Preservica Cloud Edition* (un service de préservation numérique basé sur le *cloud*), l'infrastructure technique de l'Archive, ainsi que l'analyse de rentabilité et le financement du projet pilote. Il se termine par les principaux enseignements tirés et les projets futurs. Janvier 2015 (9 pages).

Étude de cas des Archives nationales du Royaume-Uni : archives parlementaires
http://www.nationalarchives.gov.uk/documents/Cloud-Storage-casestudy_Parliament_2015.pdf
Cette étude de cas couvre les Archives parlementaires et leur expérience de l'approvisionnement via le cadre G-Cloud. Pour plus de résilience/une stratégie de sortie, elles ont choisi deux fournisseurs de services *cloud* avec différentes infrastructures de stockage sous-jacentes. Il s'agit d'un exemple d'Archive utilisant un ensemble hybride de solutions de stockage (en partie dans le *cloud* public et en partie installé localement) pour la préservation numérique, car l'Archive dispose d'un système de préservation installé localement (*Preservica Enterprise Edition*) qui est intégré au stockage dans le *cloud* et au stockage local et stocke les documents sensibles localement, et non dans le *cloud*. Janvier 2015 (6 pages).

Étude de cas des Archives nationales du Royaume-Uni : *Bodleian Library*, université d'Oxford

http://www.nationalarchives.gov.uk/documents/Cloud-storage-casestudy_Oxford_2015.pdf

Cette étude de cas porte sur la *Bodleian Library* et l'Université d'Oxford, et sur la mise en place d'une infrastructure locale « en *cloud* privé » pour ses collections numériques comprenant des livres, des images et du multimédia numérisés, des données de recherche et des catalogues. Elle explique le contexte organisationnel, la nature de ses exigences et approches en matière de préservation numérique, ses services de stockage, son infrastructure technique, ainsi que l'analyse de rentabilité et le financement. Il se termine par les principaux enseignements tirés et les projets futurs. Janvier 2015 (6 pages).

Projet Kindura du King's College London

http://link.springer.com/article/10.1186%2F2192-113X-2-13

Le projet Kindura, dirigé par le *King's College London* et financé par Jisc, visait à piloter l'utilisation d'un *cloud* hybride pour la gestion des données de recherche. Il a utilisé DuraCloud pour servir d'intermédiaire entre les ressources de stockage ou de calcul fournies par des services *cloud* externes, des services partagés ou des services internes. Il existe une étude de cas antérieure préparée par Jisc, mais un article en libre accès plus récent sur le projet est lié.

Évaluation 2011 d'Archivematica par l'Université de l'Illinois

http://e-records.chrisprom.com/evaluating-open-source-digital-preservation-systems-a-case-study-2/

Angela Jordan décrit une évaluation réalisée en 2011 par les archives de l'université de l'Illinois sur Archivematica, un système de préservation numérique en Open Source conforme au modèle de référence OAIS. Comme Archivematica était alors dans sa phase alpha, le travail avec ce système a permis d'explorer ce qu'il offrait par rapport aux besoins des archives de l'université, et de fournir des informations aux développeurs alors qu'ils continuaient à affiner le logiciel en vue de sa mise en production.

Criminalistique numérique

Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Introduction

Dans l'esprit de beaucoup de gens, la criminalistique numérique est principalement associée à l'enquête sur des actes répréhensibles. Cependant, elle est également apparue ces dernières années comme une source prometteuse d'outils et d'approches pour faciliter la conservation et la préservation des données numériques, en particulier pour protéger et enquêter sur les preuves du passé.

Les dépositaires institutionnels et les professionnels responsables d'archives personnelles et d'autres collections numériques peuvent tirer profit de l'approche de la police scientifique pour traiter de l'authenticité, de la responsabilité et de l'accessibilité numériques. Les informations personnelles numériques doivent être traitées avec la sensibilité et la sécurité requises tout en protégeant de manière évidente leur valeur probante.

La technologie de la police scientifique permet d'identifier les problèmes liés à la protection de la vie privée, d'établir une chaîne de possession pour la provenance, d'utiliser la protection de l'écriture pour la capture et le transfert et de détecter les contrefaçons ou les manipulations. Elle peut extraire et exploiter les métadonnées et le contenu pertinents, permettre une indexation et une recherche efficaces par les professionnels de la conservation et faciliter le contrôle d'audit et la granularité des autorisations d'accès. Les progrès en cours promettent une automatisation de plus en plus efficace dans le traitement de volumes toujours plus importants d'informations numériques personnelles. Si les bonnes politiques sont mises en place, l'utilisation judicieuse des technologies de police scientifique continuera à offrir des modèles théoriques, des solutions pratiques et des aperçus analytiques.

Criminalistique en pratique

Il existe trois principes fondamentaux et essentiels en matière de criminalistique numérique : les preuves sont acquises sans être altérées ; il est possible de le démontrer ; et l'analyse est menée de manière responsable et reproductible. Les processus, le matériel et les logiciels de la police scientifique numérique ont été conçus pour garantir le respect de ces exigences.

L'assurance de l'information est essentielle. Les bloqueurs d'écriture garantissent que les informations sont saisies sans les altérer, tandis que les chaînes de préservation en termes de

traitement des preuves, de contrôle des processus, d'audit des informations, de signatures numériques et de filigranes protègent les preuves historiques contre toute altération future et toute incertitude quant à leur provenance.

L'édition sélective, l'anonymisation et le cryptage, le confinement des logiciels malveillants dans des bacs à sable et d'autres mécanismes de sécurité et de contrôle précis sont nécessaires pour garantir la protection totale de la vie privée et éviter les fuites d'informations par inadvertance. Les ordinateurs familiaux, les appareils portables et les services *cloud* partageables recèlent tous une quantité considérable d'informations personnelles et posent donc des problèmes de respect de la vie privée. Les archivistes numériques et la police scientifique partagent la nécessité de traiter les informations personnelles qui en découlent de manière responsable.

L'accent mis actuellement sur l'automatisation de l'investigation scientifique numérique est particulièrement important pour la préservation du patrimoine culturel, où cette capacité est de plus en plus essentielle dans un univers numérique qui continue à se développer de manière exponentielle. Les recherches actuelles visent à traiter de grands volumes de manière efficace et efficiente en utilisant diverses techniques d'analyse. Le traitement parallèle, par exemple, grâce à des unités de traitement graphique (GPU) spécialement conçues à cet effet, et le calcul haute performance peuvent aider les activités nécessitant un processeur important, telles que la recherche et l'indexation complètes, le filtrage et le hachage, la suppression sécurisée, l'extraction, la fusion et la visualisation.

Pour la conservation et la préservation des données numériques, il convient de souligner la manière dont l'analyse numérique attire l'attention sur l'ensemble du support numérique, généralement l'image disque, c'est-à-dire le fichier qui représente tout ce qui se trouve sur le disque original.

Technologies de criminalistique

Les technologies de police scientifique varient considérablement en termes de capacité, de coût et de complexité. Certains équipements sont coûteux, mais d'autres sont gratuits. Certaines techniques sont très simples à utiliser, d'autres doivent être appliquées avec beaucoup de soin et de sophistication. Le Consortium BitCurator a été un moment important qui a rassemblé une communauté d'utilisateurs d'outils numériques Open Source d'investigation scientifique dans le contexte des archives (Lee et al, 2014). Il existe un ensemble de plus en plus riche d'outils Open Source de police scientifique qui sont gratuits à obtenir et à utiliser - le plus important pour les archivistes, BitCurator. Ces outils constituent une très bonne introduction aux tenants et aboutissants de la criminalistique numérique, et peuvent être utilisés pour comparer et recouper les résultats d'outils commerciaux ou d'autres outils open source.

Les archivistes numériques et les spécialistes de la criminalistique partagent un besoin commun de surveiller et de comprendre comment la technologie est utilisée pour créer, stocker et gérer les informations numériques. En outre, il existe un besoin mutuel de gérer ces informations de manière responsable, conformément aux normes et aux meilleures pratiques en vigueur. Les nouvelles techniques de police scientifique favorisent le traitement des informations numériques provenant d'appareils mobiles, de réseaux, de données en direct sur des ordinateurs distants, de médias flash, de machines virtuelles, de services *cloud* et de sources cryptées. L'utilisation du cryptage commence à présenter des défis importants pour la préservation numérique. Il ne s'agit pas seulement de décrypter, mais aussi d'identifier le cryptage en premier lieu. La police scientifique numérique offre quelques solutions.

La méthodologie de police scientifique et d'archivage doit conserver la capacité à la fois d'interpréter rétrospectivement les événements représentés sur les dispositifs numériques et de réagir rapidement à l'évolution du paysage numérique par la mise en place rapide de politiques, de procédures et d'installations certifiables et responsables. Le rythme du changement a également des implications pour la formation continue des professionnels de la conservation et des archivistes, et il existe des cours de criminalistique numérique approuvés par les institutions d'archivage, d'enseignement et de préservation.

Conclusion

En conclusion, le patrimoine culturel et les archives sont confrontés à de grands défis, mais la perspective d'investigation scientifique est sans aucun doute l'une des sources les plus prometteuses de connaissances et de solutions. De même, la criminalistique numérique peut bénéficier des progrès réalisés en matière de conservation et de préservation des informations numériques.

Ce bref aperçu est basé sur de courts extraits du rapport de la *Digital Preservation Technology Watch* sur l'expertise de la criminalistique et la préservation numériques (John, 2012), avec des informations supplémentaires aimablement fournies par Jeremy Leighton John, l'auteur du rapport. Voir Ressources et études de cas pour plus de détails et d'exemples.

Ressources

Rapport de veille technologique de la DPC sur la criminalistique numérique et la préservation

http://dx.doi.org/10.7207/twr12-03

Ce rapport de la DPC rédigé en 2012 donne un large aperçu de la criminalistique numérique, avec quelques indications sur les ressources et les outils qui peuvent bénéficier au patrimoine culturel et plus particulièrement à la préservation des archives numériques personnelles (60 pages).

La criminalistique numérique et le contenu numérique naissant dans les collections du patrimoine culturel

https://clir.wordpress.clir.org/wp-content/uploads/sites/6/pub149.pdf

Ce rapport de la CLIR introduit le domaine de la criminalistique numérique dans le secteur du patrimoine culturel et explore certains points de convergence entre les intérêts des personnes chargées de la collecte et de la préservation des matériaux du patrimoine culturel nativement numérique et ceux des personnes chargées de la collecte et de la préservation des preuves juridiques (93 pages).

Archivematica
https://www.archivematica.org/wiki/Main_Page
Archivematica est un système de préservation numérique Open Source et a abordé le versement d'images disques issues de techniques de la police scientifique dans le cadre de ses processus de travail et de son ensemble d'outils.

BitCurator
https://bitcurator.net/
Le site web donne accès à des informations sur le *Consortium BitCurator* (BCC), les projets et les outils. BitCurator a développé, conditionné et documenté des outils Open Source de criminalistique numérique pour permettre aux bibliothèques, aux archives et aux musées d'extraire des objets numériques à partir de supports amovibles de manière à refléter les métadonnées et à garantir l'intégrité des objets, permettant aux utilisateurs de donner un sens aux objets et de comprendre leur contexte, et empêchant la divulgation accidentelle de données sensibles. Le consortium est une association indépendante, dirigée par la communauté des membres, qui sert d'hôte et de centre de soutien administratif, des utilisateurs et de la communauté pour l'environnement BitCurator.

Forensics wiki
https://forensicswiki.xyz/page/Main_Page
Le *Forensics Wiki* est un wiki sous licence *Creative Commons* consacré à l'information sur la criminalistique numérique. Il répertorie plus de 700 pages consacrées aux outils et techniques utilisés par les enquêteurs, aux documents et rapports importants, aux personnes et organisations concernées.

La photographie invisible - Partie 2 : Piégée : Les expériences d'Andy Warhol sur l'Amiga
http://www.nowseethis.org/invisiblephoto/posts/108[44]
Une équipe d'informaticiens, d'archivistes, d'artistes et de conservateurs ont fait équipe pour retrouver les œuvres numériques perdues d'Andy Warhol sur un ordinateur Amiga Commodore vieux de 30 ans (18 mn 52 s)

[44]NdT : Le document n'était plus accessible en ligne à la date de la présente traduction, une archive est disponible sur :https://web.archive.org/web/20150925004156/http://www.nowseethis.org/invisiblephoto/posts/108

La photographie invisible - Partie 3 : Extraterrestre : Le projet de récupération d'images de l'orbiteur lunaire
http://www.nowseethis.org/invisiblephoto/posts/384[45]
Comment les « techno-archéologues » du projet de récupération d'images de l'orbiteur lunaire ont récupéré numériquement les premières photographies de la lune prises par un ensemble de sondes spatiales non habitées dans les années 1960. (22 minutes 07 secondes)

Études de cas

Projet de courrier électronique Carcanet
https://www.manchester.ac.uk/discover/news/carcanet-press-email-preservation-project/
Un projet financé par le Jisc qui s'est attaqué au défi de la capture et de la préservation des archives de courrier électronique de *Carcanet Press*, l'une des principales maisons d'édition de poésie du http. Il a remporté le prix DPC Preservation Wward 2014 pour la sauvegarde du patrimoine numérique. Le projet a exploré et adopté plusieurs outils de découverte et d'analyse de la police scientifique, notamment le *Forensic Toolkit* (FTK) d'AccessData, l'Email Examiner de Paraben et l'eDiscovery Aid4Mail de Fookes Software.

Références

John, J. L., 2012. *Digital Forensics and Preservation*. *DPC Technology Watch Report* 12-03 November 2012. Disponible à l'adresse suivante : 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. Disponible à l'adresse suivante : https://bitcurator.net/files/2018/08/code-to-community.pdf

[45]NdT : Le document n'était plus accessible en ligne à la date de la présente traduction, une archive est disponible sur :https://web.archive.org/web/20150502084728/http://www.nowseethis.org/invisiblephoto/posts/384

Identifiants pérennes

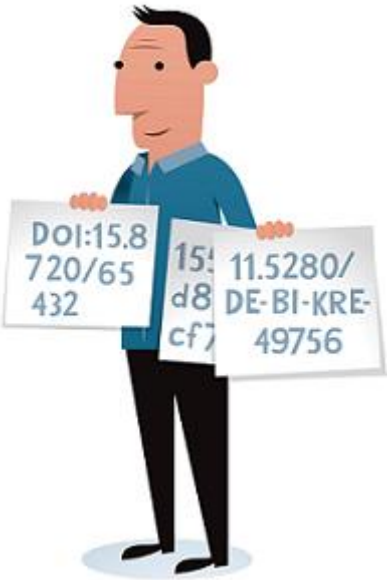


Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Introduction

Cette section fournit des conseils sur l'utilisation des identifiants pérennes pour les objets numériques et la préservation numérique. Il existe d'autres types de systèmes d'identifiants pérennes, par exemple pour les particuliers ou les institutions.

Un identifiant pérenne est une référence durable à une ressource numérique. Il se compose généralement de deux éléments : un identifiant unique et un service qui localise la ressource dans le temps, même si son emplacement change. Le premier contribue à garantir la provenance d'une ressource numérique (qu'elle est bien ce qu'elle prétend être), tandis que le second permet de s'assurer que l'identifiant pointe vers un emplacement correct.

Les identifiants pérennes visent donc à résoudre le problème de la persistance de l'accès à la ressource citée, en particulier dans la littérature universitaire. Trop souvent, les adresses web (liens) ne parviennent pas à vous conduire à la ressource référencée que vous attendez. Cela peut être dû à des raisons technologiques, comme une défaillance du serveur, mais les défaillances d'origine humaine sont plus fréquentes. Les organisations transfèrent des revues à de nouveaux éditeurs, réorganisent leurs sites web ou se désintéressent des contenus plus anciens, ce qui entraîne des liens rompus lorsque vous essayez d'accéder à une ressource. Cette situation est frustrante pour les utilisateurs, mais les conséquences peuvent être graves si la ressource liée est essentielle pour des raisons juridiques, médicales ou scientifiques.

Les identifiants pérennes peuvent également être utilisés « en coulisses » dans un entrepôt d'archives pour gérer certains des défis liés au catalogage et à la description, ou pour assurer le contrôle intellectuel et l'accès aux documents numériques natifs.

Schémas

Étant donné que le problème de la pérennité d'un identifiant est créé par les humains, la solution des identifiants pérennes doit également impliquer des personnes et des services et pas seulement des technologies. Il existe plusieurs schémas d'identificateurs pérennes et tous

nécessitent une composante humaine pour maintenir leurs systèmes de résolution. Les principaux systèmes d'identification pérenne actuellement utilisés sont détaillés ci-dessous.

Digital Object Identifier (DOI)

Les DOI sont des identifiants numériques pour les objets (qu'ils soient numériques, physiques ou abstraits) qui peuvent être attribués par les organisations membres d'une des agences d'enregistrement des DOI ; les deux plus connus sont CrossRef, pour les articles de journaux et certaines autres publications scientifiques, et DataCite pour un large éventail d'objets de données. En plus de l'identifiant de l'objet, le DOI dispose d'une infrastructure de système permettant de s'assurer qu'une URL se résout à l'emplacement correct de cet objet.

Handle

Les identifiants Handle sont des identifiants uniques et permanents pour les ressources Internet, avec un registre central pour résoudre les URL vers l'emplacement actuel. Chaque Handle identifie une ressource unique et l'organisation qui l'a créée ou qui la gère actuellement. Le système Handle détermine également l'infrastructure technique des DOI, qui sont un type spécial de Handle.

Archival Resource Key (ARK)

ARK est un système d'identification conçu par la *California Digital Library* (CDL), visant à identifier les objets de manière permanente. Le schéma a été conçu sur l'idée que la pérennité « est purement une question de service et n'est ni inhérente à un objet ni conférée à celui-ci par une syntaxe de nommage particulière. »

Persistent Uniform Resource Locator (PURL)

Les PURL sont des URL qui redirigent vers l'emplacement de la ressource web demandée en utilisant des codes d'état HTTP standard. Un PURL est donc une adresse web permanente qui contient la commande de redirection vers une autre page, qui peut changer au fil du temps

Universal Resource Name (URN)

Les URN sont des identifiants pérennes, indépendants du lieu, qui permettent de cartographier simplement les espaces de noms dans un seul espace de noms URN. L'existence d'un tel identificateur de ressource uniforme n'implique pas la disponibilité de la ressource identifiée, mais ces URI doivent rester uniques et persistants au niveau mondial, même lorsque la ressource cesse d'exister ou devient indisponible. Le terme URN est désormais déprécié, sauf au sens très étroit d'un espace de noms formel pour exprimer un *Uniform Resource Identifier*.

Choisir un système d'identification pérenne

Il faut un écosystème pour maintenir la pérennité du service de résolution - soit par l'organisation qui héberge la ressource numérique, soit par un tiers de confiance, soit par une combinaison des deux. Chaque schéma a ses propres avantages et contraintes, mais il est utile de tenir compte des éléments suivants pour décider d'une stratégie ou d'une approche d'identification pérenne :

Avantages

- Il est d'une importance capitale pour aider à établir l'authenticité d'une ressource.
- Permet d'accéder à une ressource même si son emplacement change.
- Permet de surmonter les problèmes causés par la nature instable des URL.
- Permet l'interopérabilité entre les collections.

Inconvénients

- Il n'existe pas de système unique accepté par tous, bien que les DOI soient très bien établies et largement déployées.
- L'établissement ou l'utilisation d'un service de résolution peut entraîner des coûts.
- Dépendance à une maintenance continuelle du système d'identification pérenne.

Conclusions

Les identifiants pérennes doivent être pris en charge par des services durables et ne sont pas simplement des chaînes uniques de caractères alphanumériques qui sont attribuées à une ressource numérique. Ils sont devenus particulièrement importants pour les données de recherche et les articles de périodiques électroniques (voir la section sur la préservation du contenu spécifique des périodiques électroniques) et constituent une partie importante de l'infrastructure à long terme pour la préservation numérique de la recherche. Pour la question du lien pour les pages web plus générales et les solutions exploitant les archives web pour résoudre ce problème, voir la section sur la préservation spécifique du contenu dans l'archivage web.

Ressources

Les identifiants pérennes - un aperçu. Revue de la veille technologique TWR

http://www.metadaten-twr.org/2010/10/13/persistent-identifiers-an-overview/

Cet article de Juha Hakala (2010) décrit cinq systèmes d'identifiants pérennes (ARK, DOI, PURL, URN et XRI) et compare leurs fonctionnalités avec les « cool URIs ». L'objectif est de fournir une vue d'ensemble, et non de donner un quelconque classement de ces systèmes.

Préservation, confiance et accès continu pour les périodiques électroniques - Rapport de veille technologique de la DPC

http://dx.doi.org/10.7207/twr13-04

Ce rapport de 2013 par Neil Beagrie s'intéresse aux développements actuels et aux problèmes auxquels sont confrontés les bibliothèques, les éditeurs, les intermédiaires et les fournisseurs de services dans le domaine de la préservation numérique, de la confiance et de l'accès permanent aux périodiques électroniques. Il inclut des leçons générales et des recommandations sur l'externalisation et le maintien de la confiance pour la communauté de la préservation numérique au sens large et couvre les aspects pertinents en matière légale, économique et fonctionnelle aussi bien que technologique (49 pages).

Identifiants pérennes dans la publication et la citation de données scientifiques

https://www.researchgate.net/publication/241581975_Persistent_Identifiers_in_the_Publication_and_Citation_of_Scientific_Data_-_Theory_and_Practice

Présentation de Jens Klump, Centre allemand de recherche en géosciences (GFZ) sur le projet STD-DOI du DFG, qui détaille le contexte et le raisonnement derrière la fondation de DataCite 2009 (47 pages).

Document d'information de la DCC : Identificateurs pérennes

http://www.dcc.ac.uk/resources/briefing-papers/introduction-curation/persistent-identifiers

Ce document de Joy Davidson, publié en 2006, examine comment les progrès réalisés dans la définition de la nature et des exigences fonctionnelles des systèmes d'identification sont entravés par l'absence d'accord commun sur ce que les identificateurs doivent réellement faire : simplement fournir un nom unique au niveau mondial ou local pour une ressource numérique ou analogique, ou intégrer des services associés tels que la résolution et la liaison de métadonnées. L'application et la maintenance des identificateurs ne constituent qu'une partie d'une stratégie globale de préservation numérique ; pour offrir des garanties de pérennité à long ou à court terme, ils nécessitent un engagement institutionnel et des rôles et responsabilités clairement définis(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
http://archive.org/services/purl/
URN
http://tools.ietf.org/html/rfc3986

Études de cas

Étude de cas de la DCC : Attribution d'identifiants d'objets numériques aux données de recherche à l'Université de Bristol
http://www.dcc.ac.uk/resources/persistent-identifiers
L'Université de Bristol gère un dépôt de données de recherche dédié dans le cadre de son service de données de recherche. Ils utilisent le service DataCite de la *British Library* pour attribuer des identificateurs d'objets numériques (DOI) aux ensembles de données de recherche afin de fournir des identificateurs uniques et perpétuels pour les données, afin de permettre une citation et une découverte faciles. Le service de données de recherche de Bristol fournit des conseils sur la manière d'utiliser les identificateurs pour citer les données et développe des politiques appropriées pour surveiller l'utilisation. 2004. (4 pages).

Des liens qui durent

http://www.dpconline.org/events/previous-events/925-links-that-last

Cette journée d'information de la DPC en juillet 2012 a permis de présenter les thèmes des identifiants pérennes et des données liées, et de discuter des implications pratiques des deux approches de la préservation numérique. Elle a examiné la viabilité des services qui offrent des identifiants pérennes et ce qu'ils offrent dans le contexte de la préservation ; elle a passé en revue les récents développements en matière de données liées, en considérant comment ces ensembles de données pourraient être préservés ; et en introduisant ces deux sujets parallèles, elle a ensuite examiné si les deux approches peuvent être liées de manière à créer une nouvelle classe de données liées robustes. Une série de présentations comprenant des études de cas sont reliées à partir du programme provisoire.

Préservation du contenu spécifique

Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

A qui s'adresse ce chapitre ?

Encadrants (DigCurV : *Manager Lens*) et personnel opérationnel (DigCurV : *Practitioner Lens*), éditeurs et autres créateurs de données, fournisseurs de services tiers.

Niveau de connaissance présumé

Novice à intermédiaire.

Objectifs

- Établir une connexion et une résonnance avec les rapports de la série *DPC Tech Watch*. Ces rapports fournissent des «plongées approfondies» de niveau avancé dans des domaines spécifiques de la préservation du contenu (par exemple, les e-mails) qui peuvent être cités ou utilisés pour fournir des études de cas dans ce manuel.
- Être développé par la DPC par le biais de mises à jour et d'ajouts à la série *Tech Watch* afin d'en faciliter la maintenance, la rentabilité et la durabilité à long terme.
- Fournir un bref aperçu et des études de cas, adaptés aux utilisateurs novices ou de niveau intermédiaire sur les questions de préservation numérique de types de contenu spécifiques couverts par les rapports de veille technologique de la DPC. Trois types de contenu sont actuellement couverts : les périodiques électroniques, les images animées et les sons et l'archivage du web. Nous espérons en ajouter d'autres ultérieurement.

Périodiques électroniques

Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Aperçu

Cette étude de cas donne un bref aperçu de la question de la préservation des périodiques électroniques (niveau débutant à intermédiaire) et constitue un résumé, mis à jour et augmenté par l'auteur du rapport de veille technologique de la DPC intitulé *Preservation, Trust and Continuing Access for e-Journals*. Deux « mini études de cas » sont incluses, ainsi que de brefs résumés des principaux services et solutions. Le rapport lui-même est recommandé aux lecteurs qui ont besoin d'aller plus loin sur le sujet et sa mise en œuvre. Il couvre un plus large éventail de questions et de pratiques de manière plus approfondie, avec des lectures et des conseils supplémentaires (Beagrie, 2013).

Introduction

La préservation des contenus et la préservation de l'accès au contenu numérique sont devenues de plus en plus importantes pour les bibliothèques de recherche à mesure que les périodiques et les articles publiés sont passés du format imprimé au format électronique. Les modèles commerciaux, éditoriaux et de distribution traditionnels ont également subi des transformations majeures à la suite de ce changement.

Parmi les nombreux changements importants, on peut citer le passage des bibliothèques qui achetaient, détenaient et conservaient physiquement un périodique papier dans leurs collections locales (avec de multiples redondances d'exemplaires entre les bibliothèques), à la location (licence) d'un accès à distance à un périodique électronique détenu sur des plateformes d'éditeurs souvent basées à l'étranger et donc soumises à d'autres réglementations.

Parallèlement, on observe un mouvement croissant de libre accès aux articles de périodiques électroniques qui cherche à supprimer les frais d'abonnement. Les périodiques par abonnement, les périodiques en libre accès et les hybrides des deux (soit un mélange d'articles en libre accès et par abonnement dans un périodique, soit une « barrière mobile » qui permet un libre accès après une période déterminée) offrent un paysage complexe pour la préservation des périodiques électroniques et leur accès à long terme.

Le paysage des périodiques électroniques continue d'évoluer à mesure que l'édition électronique elle-même commence à passer du contenu statique au contenu dynamique, et que l'importance des données et des objets supplémentaires liés aux articles augmente dans les principales disciplines.

Tous ces changements ont peu à peu rendu la préservation des périodiques électroniques plus exigeante, plus internationale et plus interdépendante, et ont mis en avant les questions de confiance dans les services de préservation et de consultation. Dans ce contexte, la confiance ne concerne pas seulement la technologie de préservation, mais aussi la négociation des droits, et la préservation d'une trace de ces droits pour un usage futur, ainsi que l'obtention d'informations transparentes sur ce qui est archivé, comment il est préservé, et comment et quand il peut être consulté.

Cela fait des périodiques électroniques l'un des domaines les plus dynamiques et les plus difficiles de la préservation numérique, notamment en termes de modèles commerciaux et de mécanismes de confiance pour les services de préservation partagés ou externalisés.

Services et solutions

Il est important de comprendre les implications significatives pour la préservation et l'accès qui s'appliquent aux périodiques électroniques : en particulier la distinction entre l'accès continu et la préservation à long terme, car ces différences conduisent à différents types de services pour l'archivage des périodiques électroniques.

- L'accès continu (également dénommé accès garanti après l'abonnement ou accès perpétuel) s'applique uniquement aux périodiques par abonnement et à la garantie d'un accès à long terme pour leurs abonnés ;
- La préservation à long terme s'applique à la fois au contenu ouvert et au contenu auquel l'abonné a souscrit.

Les principaux services et solutions de préservation et d'accès continu disponibles pour les périodiques électroniques sont les suivants :

Keepers Registry
Le *Keepers Registry* est un service du *Joint Information Systems Committee* (Jisc) qui fournit des informations facilement accessibles sur l'inclusion des périodiques électroniques dans les services de préservation et met en évidence les périodiques électroniques pour lesquelles il n'existe pas de dispositions d'archivage. EDINA, un centre de données national basé à l'Université d'Edimbourg, a développé ce service avec son partenaire dans le projet, le Centre international de l'ISSN à Paris.
Dépôt légal et volontaire dans les bibliothèques de droits d'auteur
Le rôle d'une bibliothèque nationale est de veiller à ce que le patrimoine publié de son pays soit conservé et accessible. Dans de nombreux pays, le dépôt légal est l'un des moyens pour y parvenir. La tendance mondiale est à l'extension du dépôt légal de l'imprimé aux publications numériques (périodiques, livres électroniques). La législation sur le dépôt légal (ou des dispositions similaires de dépôt volontaire) concerne normalement les périodiques électroniques par abonnement considérés comme faisant partie du patrimoine national publié de ce pays. Afin de protéger les intérêts commerciaux de l'éditeur, elle restreint également l'accès hors site à l'objet électronique préservé pendant au moins la période de protection de la propriété intellectuelle (licences, droits

d'auteurs et droits voisins du droit d'auteur). Cela signifie généralement qu'une collection nationale de dépôt légal ne couvre pas la gamme internationale de périodiques électroniques par abonnement sous licence d'autres bibliothèques et de leurs utilisateurs, et ne répond pas à leurs exigences en matière de droits d'accès continu.

CLOCKSS

CLOCKSS (*Controlled LOCKSS*) est une association à but non lucratif entre les bibliothèques et les éditeurs. Il s'agit d'un coffre-fort, basé sur le logiciel LOCKSS (voir la section ci-dessous sur LOCKSS) dans lequel un nombre limité de bibliothèques assument un rôle d'archivage au nom d'une communauté plus large. Il fournit aux bibliothèques l'assurance que les périodiques électroniques et les autres contenus auxquels elles ont souscrit seront préservés à long terme. Il est décrit comme un « réseau LOCKSS privé ».

e-Dépôt de la KB

La *Koninklijke Bibliotheek* (KB) est la Bibliothèque nationale des Pays-Bas et gère l'e-Depot. Elle a pris la décision politique d'archiver les périodiques qui relèvent de son mandat national et, en outre, une série de périodiques électroniques (y compris les titres en libre accès dans le Répertoire des périodiques en libre accès) publiés au-delà de ses frontières. L'e-Depot ne prévoit pas actuellement l'accès continu après suppression du contenu par les détenteurs de licences. En général, l'accès des utilisateurs finaux est limité à la consultation sur place au KB pour des raisons de recherche privée uniquement et l'accès en ligne est refusé. Toutefois, l'accès en ligne complet est accordé aux publications des éditeurs à accès libre.

LOCKSS

LOCKSS (*Lots of Copies Keep Stuff Safe*) fournit aux bibliothèques des outils et un soutien en matière de logiciels libres afin qu'elles puissent assurer la garde locale d'une grande variété de documents, y compris les abonnements et les actifs scientifiques en libre accès (livres, périodiques, etc.). Les lecteurs accèdent au contenu préservé par LOCKSS lorsque (et pour quelque raison que ce soit) l'objet ne peut être consulté sur les serveurs de l'éditeur (ou de l'intermédiaire). La nature hautement distribuée de cette approche vise à garantir une réplication suffisante pour sauvegarder le contenu malgré les catastrophes potentielles qui pourraient frapper les institutions LOCKSS individuelles.

Portico

Portico est conçu spécifiquement comme un service tiers pour la littérature scientifique publiée sous forme électronique et fournit trois services de préservation spécifiques pour les périodiques électroniques, les livres électroniques et les collections historiques numérisées respectivement. Il fournit aux bibliothèques l'assurance que les périodiques électroniques et les autres contenus auxquels elles ont souscrit seront préservés à long terme. Portico ne donne accès aux périodiques électroniques qu'après des « événements déclencheurs » spécifiques. En outre, si un éditeur a désigné Portico comme tel, il peut également servir de mécanisme potentiel d'accès post-annulation.

Hébergement en consortium

Un petit nombre de consortiums régionaux organisent et fournissent également leurs propres services d'hébergement pour l'accès et la préservation des périodiques électroniques. Parmi les exemples notables, citons OhioLink, géré par l'*Ohio Library and Information Network*, et le *Scholars Portal*, géré par l'*Ontario Council of University Libraries*.

Étude de cas n° 1 : le périodique électronique ou ses anciens numéros ne sont plus disponibles auprès de l'éditeur

C'est un scénario très probable lorsque que les éditeurs fusionnent ou modifient leurs modèles commerciaux, lorsque les grands éditeurs revoient et adaptent leur portefeuille de titres, ou encore lorsque les sociétés savantes transfèrent les contrats de publication de leurs revues d'un éditeur à l'autre. Les titres des revues sont aussi parfois échangés entre éditeurs, ce qui peut signifier que l'accès aux numéros précédents n'est plus pris en charge par le propriétaire précédent.

L'initiative *Transfer Code of Practice* de l'UKSG (dont fait partie Couperin) a produit un code de bonnes pratiques visant à atténuer les problèmes créés lorsque les titres de revues passent d'un éditeur à l'autre. Les paragraphes suivants, contenus dans la version 3 du code (UKSG, 2014), présentent un intérêt particulier :

- L'éditeur cédant informera l'éditeur destinataire de toutes les dispositions de préservation existantes pour la revue.
- L'éditeur cédant doit garantir un accès continu à ses abonnés lorsqu'il a accordé des droits d'accès perpétuels, même si l'éditeur cédant cesse d'héberger la version en ligne de la revue après la date de transfert effective.
- L'éditeur cédant ou l'éditeur récepteur, ou les deux, peuvent remplir les obligations d'accès perpétuel. À dessein, le code ne précise pas les moyens d'obtenir cet accès, mais confie à l'éditeur cédant la responsabilité de s'assurer que les abonnés auxquels il a accordé des droits d'accès perpétuels continueront à avoir accès après le transfert.
- L'éditeur cédant fera des efforts raisonnables pour communiquer les informations relatives au transfert des revues lorsque des droits d'accès perpétuels ont été accordés dans le cadre d'un accord de licence, à moins que les droits d'archivage ne restent chez l'éditeur cédant.
- Les abonnés qui se sont vu accorder des droits d'accès perpétuels à un contenu précédemment publié avec l'autorisation du propriétaire de la revue doivent voir ces droits honorés.
- L'éditeur cédant ou l'éditeur destinataire, ou les deux, peuvent remplir les obligations d'accès perpétuel.
- L'éditeur destinataire maintiendra les dispositions de préservation existantes, ou équivalentes, pour la revue après la date de transfert effective.
- L'éditeur destinataire ne retirera pas le contenu qui a été précédemment déposé dans un entrepôt numérique de préservation, même si l'éditeur destinataire ne continue pas à déposer le contenu dans l'entrepôt.

La décision de l'éditeur Sage de ne plus proposer sa publication Graft a fourni un exemple concret d'accès déclenché à partir de trois solutions d'archivage - Portico, KB e-Depot et CLOCKSS. Dans ce cas, toutes ont pu continuer à offrir l'accès aux numéros qu'elles détenaient, soit en accès libre (CLOCKSS et KB e-Depot), soit en tant que service aux membres (Portico). Bien qu'il ne soit pas possible de garantir que les archives comprendront tous les anciens numéros du titre (comme dans le cas de Graft), la participation à une solution d'archivage qui couvre au moins certains numéros réduira considérablement le risque d'interruption du service.

Étude de cas n° 2 : périodiques électroniques de la bibliothèque, accès perpétuel et sortie définitive des imprimés

Cette étude de cas a été publiée pour la première fois par Jisc dans le cadre des travaux financés par son programme de préservation numérique et a été intégrée au rapport Tech Watch. Elle a été adaptée pour être utilisée dans le manuel.

L'étude de cas se distingue des autres en illustrant quelques problèmes liés à la réalisation de certaines des économies potentielles des périodiques électroniques, notamment en termes d'espace. De plus en plus, les bibliothèques universitaires investissent massivement dans des périodiques électroniques qui font double emploi avec leurs anciens tirages. Pour les bibliothèques confrontées à une forte pression sur l'espace, une solution à leur problème consiste à éliminer les anciens tirages qui chevauchent leurs fonds électroniques.

L'étude de cas se concentre sur le travail de la bibliothèque de l'*Imperial College London Library*, qui fournit une base de données et une boîte à outils au personnel chargé de prendre les décisions de désherbage (Cooper et Norris, 2007). L'*Imperial College* a ainsi établi trois critères pour déterminer la viabilité des périodiques électroniques par rapport aux imprimés. Leur accès a été classé comme durable lorsqu'au moins l'un des critères suivants s'appliquait :

- L'*Imperial College* avait des droits d'accès perpétuels au contenu, via le web. Les droits d'accès perpétuels de l'*Imperial College* étaient loin d'être aussi complets qu'ils l'auraient souhaité ; ils estimaient que moins de 50 % de leur contenu était couvert. En outre, certaines de leurs licences spécifiaient une méthode de diffusion inappropriée pour l'accès après la fin de la licence. Comme ils ne prenaient plus en charge les CD-ROM sur le réseau et ne disposaient pas des ressources nécessaires pour monter le contenu des revues sur leurs réseaux locaux, ils considéraient qu'une revue n'était viable que si l'accès perpétuel était fourni via le web.
- La revue était en accès libre permanent pour toutes les années ou certaines années. Les périodiques hybrides en accès libre n'ont pas été inclus dans cette catégorie, car le projet ne s'intéressait pas à la durabilité au niveau des articles. Trouver des revues en libre accès qui remplissaient leurs critères s'est avéré plus difficile que prévu. La principale pierre d'achoppement était leur besoin d'assurance sur la permanence de l'accès. Bien que les déclarations de Bethesda et de Berlin sur le libre accès incluent l'accès perpétuel dans leurs définitions, l'*Imperial College* a découvert que toutes les revues en libre accès ne répondaient pas à ce critère de permanence.
- Le contenu se trouvait dans un des services de confiance de l'*Imperial College* comme JSTOR, les archives numériques d'ACM ou une Archive financée par le Jisc. L'*Imperial College* a noté que parmi ses trois critères de durabilité, celui-ci, couvrant les services qui n'offrent pas de droits d'accès perpétuels, était le plus difficile à cerner. Les services qui entrent dans cette catégorie partagent tous deux caractéristiques : le premier est un bon historique de stabilité, c'est-à-dire qu'ils ont démontré la continuité des titres d'une année à l'autre pendant toute la durée de leur abonnement ; le second est un historique et une réputation, un caractère abordable et un bon rapport qualité-prix.

Vingt et un mois après le début du projet, l'*Imperial College* avait identifié 700 mètres linéaires à désherber sur ses sites et prévoyait de déployer le chantier de désherbage sur d'autres. Bien qu'il n'en soit qu'au début, les critères de durabilité semblent fonctionner. Le seul contenu durable perdu était quatre revues du même éditeur, et ils contestent actuellement cette perte. La

documentation des droits acquis et des licences dans une base de données a été véritablement bénéfique; sans elle, ils n'auraient pas eu conscience de la perte d'un contenu sur lequel ils avaient pourtant acquis des droits d'accès perpétuels.

Conclusions

L'accès perpétuel et la préservation des périodiques électroniques ont impliqué des initiatives visant à organiser une collaboration inter-institutionnelle, à développer des services tiers et à définir une confiance dans l'accès et la préservation à long terme entre les différentes parties prenantes. Les questions qu'elle a dû traiter vont bien au-delà de la technologie. Les questions juridiques, économiques et de services sont tout aussi essentielles à son succès. De nombreux défis restent à relever dans le domaine de l'archivage des périodiques électroniques, mais certains succès importants et leçons apprises intéressent l'ensemble de la communauté de la préservation numérique ainsi que les bibliothèques et les éditeurs.

Ressources

Preservation, Trust and Continuing Access for e-Journals, DPC Technology Watch Report 13-04 septembre 2013

http://dx.doi.org/10.7207/twr13-04

Ce rapport examine les développements et les problèmes actuels auxquels sont confrontés les bibliothèques, les éditeurs, les intermédiaires et les fournisseurs de services dans le domaine de la préservation numérique, de la confiance et de l'accès continu aux périodiques électroniques. Il comprend également des enseignements et des recommandations générales sur l'externalisation et la confiance, tirés dans ce domaine qui intéresse la communauté de la préservation numérique au sens large. Il n'est pas uniquement axé sur la technologie et couvre des questions juridiques, économiques et de services pertinentes (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

De plus en plus, les bibliothèques universitaires investissent massivement dans des périodiques électroniques qui reproduisent leurs anciens tirages. Pour les bibliothèques confrontées à une forte pression sur l'espace, une solution à leur problème consiste à éliminer les anciens tirages qui chevauchent leurs fonds électroniques. Cet article de Cooper et Norris, publié en 2007, décrit le travail effectué à l'*Imperial College London Library* pour fournir un ensemble d'outils au personnel qui prend de telles décisions de désherbage.

UKSG, 2014 *Transfer Code of Practice*: Version 3.0 March 2014

https://www.uksg.org/sites/uksg.org/files/TRANSFER_Code_of_Practice_3%200_FINAL.pdf

Le Code de bonnes pratiques de transfert promeut un ensemble de normes qui s'appliquent chaque fois qu'une revue est transférée d'un éditeur ou d'une plateforme de publication à un autre. Les éditeurs qui souscrivent publiquement au code et l'appliquent dans leur pratique sont considérés comme conformes au transfert. Les participants du secteur s'engagent volontairement,

ainsi le Code de bonnes pratiques en matière de transfert ne remplace pas les conditions contractuelles, les droits de propriété intellectuelle ou le marché concurrentiel entre éditeurs.

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

Références

Beagrie, N., 2013. *Preservation, Trust and Continuing Access for e-JournalsDPC Technology Watch Report* 13-04 September 2013. Disponible à l'adresse suivante : 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. Disponible à l'adresse suivante : https://spiral.imperial.ac.uk/handle/10044/1/503

UKSG, 2014.*Transfer Code of Practice*: Version 3.0 March 2014. Disponible à l'adresse suivante : https://www.uksg.org/sites/uksg.org/files/TRANSFER_Code_of_Practice_3%200_FINAL.pdf

Images animées et son

Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Aperçu

Cette étude de cas donne un bref aperçu de la question de la préservation des images animées et du son (niveau débutant à intermédiaire) et constitue un résumé du rapport de veille technologique de la DPC intitulé *Technology Watch Report on Preserving Moving Picture and Sound*. Cinq « mini études de cas » d'institutions britanniques qui ont mené des projets de préservation et d'accès au contenu d'images animées et de sons sont incluses. Le rapport lui-même constitue une étude approfondie qui aborde un plus large éventail de questions et de pratiques, avec de nombreuses lectures et des conseils supplémentaires (Wright, 2012).

Introduction

Le domaine audiovisuel est unique en ce sens que la numérisation est quasi systématiquement essentielle à la préservation. La numérisation audiovisuelle pour la préservation est si omniprésente que les deux mots en sont venus à être utilisés de manière interchangeable. L'audio et la vidéo ont besoin de la numérisation pour la survie même de leur contenu, en raison de l'obsolescence des équipements de lecture, de la détérioration et de l'obsolescence des éléments physiques, qu'ils soient analogiques ou numériques. Le problème technologique fondamental pour les collections d'images animées et de sons est la nécessité de numériser tous les contenus qui se trouvent actuellement sur les rayonnages. Les films sur les rayonnages peuvent être conservés (sauf s'ils se détériorent déjà), mais ils doivent encore être numérisés pour être accessibles.

L'accès est une question vitale pour la préservation : il est difficile, voire impossible, de motiver et de financer la numérisation à des fins de préservation uniquement. Le public, les institutions et l'enseignement s'intéressent beaucoup à l'enregistrement audiovisuel du XXe siècle. La création d'un accès à ce type de document est la clé pour obtenir le soutien nécessaire à la numérisation et à la préservation du contenu.

Le paysage de l'image animée et du son est complexe : physiquement, il existe de grandes différences entre les enregistrements audio, vidéo et cinématographiques. Les formats et les équipements d'enregistrement/lecture sont complètement séparés ; les procédures de numérisation sont différentes ; les fichiers numériques ont des formats de conteneurs et des métadonnées différents (avec quelques chevauchements) ; et les exigences de stockage diffèrent,

la vidéo nécessitant environ 100 fois plus de stockage par seconde d'objet enregistré que l'audio, et le film numérique haute résolution nécessitant environ 10 fois plus de stockage que la vidéo.
En outre, sur le plan culturel et économique, il existe des différences importantes en matière de préservation entre les collections constituées par :

- Les industries culturelles et commerciales audiovisuelles où la préservation nécessite une justification commerciale et une analyse de rentabilité ;
- Les organismes publics – (radiodiffusion de service public, collections universitaires et institutions patrimoniales telles que les musées nationaux, les bibliothèques et les instituts cinématographiques), où la préservation doit être justifiée par la constitution d'un patrimoine culturel, bien que ce secteur ait de plus en plus besoin d'une analyse de rentabilité ;
- Les organismes de recherche et les industries de l'ingénierie tels que la médecine, la géologie et la surveillance, où les enregistrements d'images ou d'événements sismiques sont des données brutes, conservées comme dossiers médicaux ou pour retraitement ;
- Les initiatives privées de collecte - un large éventail de collections indépendantes, allant d'efforts individuels à des institutions spécialisées à but non lucratif (par exemple, des clubs de machines à vapeur ou des recherches ethnologiques) qui n'entrent dans aucune des catégories ci-dessus, bien que leurs objets pourraient finalement être donnés à une collection publique.

Dans ce paysage, on trouve toute une série de technologies, notamment l'ingénierie, l'informatique, les technologies du web et de l'internet, l'archivage, la gestion des médias, la gestion des collections de musée, la conservation, la préservation, l'accès, la gestion des connaissances et la découverte des ressources.

Défis techniques

Les enregistrements audiovisuels sont une réalité de substitution. La technologie permet à l'auditeur et au téléspectateur de se faire une idée de ce à quoi une situation ressemblait, mais la technologie ne saisit en fait que la séquence des motifs lumineux ou des pressions sonores agissant sur l'instrument d'enregistrement (caméra, microphone). Ces motifs (pour les films) et ces signaux (pour la vidéo et l'audio) ressemblent plus à des données qu'à des objets. L'exigence de préservation n'est pas de conserver le support d'enregistrement original, mais de conserver les données, les informations, récupérées sur ce support.

L'un des principaux problèmes technologiques est le transfert du contenu numérique des supports optiques et magnétiques (tels que les CD et DVD, les bandes vidéo numériques, les DAT [*Digital Audio Tape*, format de cassette audio, NdT] et les minidisques) vers des fichiers. Cette extraction de contenu de numérique à numérique est un domaine de la préservation numérique unique au monde audiovisuel, et pose des problèmes non résolus de contrôle des erreurs dans le processus d'extraction et de transfert.

Le dernier domaine technologique est la préservation numérique du contenu des fichiers qui résultent de la numérisation ou de l'extraction, et des fichiers nativement numériques. Si une grande partie de cette préservation présente des problèmes et des solutions communes avec d'autres contenus, il existe un problème spécifique de préservation de la qualité du signal numérisé qui est, là encore, propre aux contenus audiovisuels. La gestion de la qualité par des cycles de codage, de décodage et de reformatage avec perte est l'un des principaux défis de la préservation numérique des fichiers audiovisuels. L'autre problème est la gestion des métadonnées intégrées.

Depuis trois décennies pour l'audio, et depuis au moins deux décennies pour la vidéo, les services d'archives numérisent leur contenu analogique à des fins de préservation et d'accès. Les domaines qui posent problème sont les suivants :

- Une lecture réussie des originaux, afin d'obtenir un signal optimal pour la numérisation ;
- Les normes : quel niveau de compression, quelle méthode d'encodage et quel format de fichier utiliser ;
- L'efficacité : numériser les objets analogiques existants suffisamment rapidement et économiquement pour faire face à l'ampleur et à l'urgence du problème.

Étapes de la préservation numérique du son et de l'image animée

Pour la préservation du son et des images animées, les étapes suivantes du processus global doivent être clairement définies :

- Le signal : le son provenant d'un microphone, le signal vidéo provenant d'une caméra vidéo. Ces signaux ont des propriétés physiques (largeur de bande ; gamme dynamique) qui peuvent être définies et mesurées. La qualité d'un enregistrement et le succès ou l'échec de tout processus de copie, de numérisation ou de préservation peuvent être réduits (en grande partie) à la mesure dans laquelle ce processus maintient ces deux propriétés physiques du signal original ;
- L'enregistrement d'un signal sur un support (*carrier* en anglais, NdT) (également appelé support *[support* en anglais, NdT], médium physique ou format d'enregistrement). Pendant un siècle, les méthodes de capture d'un signal étaient liées au support du signal : un cylindre de cire, une bobine de film ou une bande vidéo. La technologie numérique produit des enregistrements qui sont indépendants des supports. L'indépendance du support est une libération : les disques, les bandes et les films se détériorent ou s'abîment. Les enregistrements numériques natifs sont libérés de ces problèmes liés au support, ce qui conduit à un désir de libérer les enregistrements analogiques par la numérisation ;
- La numérisation : les enregistrements analogiques peuvent être lus et enregistrés sur un nouveau support, ou numérisés et ainsi libérés de la dépendance vis-à-vis du support. La numérisation doit garantir que la version numérique a la même largeur de bande et la même gamme dynamique que l'original, afin de conserver la qualité originale
- La préservation numérique de la représentation numérique d'un signal, c'est-à-dire la préservation des nombres, mais aussi la préservation de la technologie nécessaire pour décoder (rendre) les nombres. Le contenu audiovisuel pose un problème particulier. Le codage du signal peut être un compromis, en ne capturant pas la totalité du signal, mais en en perdant une partie (codage avec perte) pour obtenir une représentation plus compacte, ce qui réduit les coûts de stockage et de transmission. Malheureusement, les codeurs/décodeurs (codecs) ne sont plus utilisés et sont remplacés par des technologies plus récentes. Le format de fichier contenant le signal codé, le *wrapper*, est également sujet à l'obsolescence. La défaillance et l'obsolescence de la technologie de stockage et l'obsolescence des méthodes de codage/décodage et des formats de *wrapper* sont des problèmes majeurs de préservation numérique du contenu audiovisuel.

Accès et droits

Les contenus sonores et cinématographiques issus du cinéma, de la radiodiffusion et de l'industrie musicale commerciale sont limités par les questions de droits. La musique est protégée par le droit d'auteur pour le compositeur et pour l'objet physique contenant une performance (droits voisins). Les productions cinématographiques sont protégées, et la musique utilisée dans un film conserve ses protections distinctes. La radiodiffusion est encore plus compliquée, car toutes les parties

impliquées dans une production peuvent avoir des droits sur l'exploitation future après une ou deux transmissions qui ont été spécifiées dans des contrats types. Ces droits sont considérés comme une protection par les titulaires de droits, mais aussi comme des restrictions d'accès. La situation est particulièrement difficile pour un radiodiffuseur public. Le public a invariablement le sentiment que toute production d'un radiodiffuseur public a déjà été payée par lui, qu'elle est déjà propriété publique et qu'elle devrait être accessible au public. Malheureusement, ce sentiment compréhensible n'est pas le même que la définition juridique régissant le moment où une œuvre entre dans le domaine public (généralement déterminé par les dates d'expiration des droits d'auteur et des droits associés).

Étude de cas n° 1 : le projet d'accès aux ressources vidéo de *Open University* (OU)

Il s'agit d'un projet d'accès et de réutilisation. L'objectif est de numériser (lorsque nécessaire) les contenus audiovisuels précédemment créés par l'OU, et de les placer dans un système de gestion des contenus afin que ces contenus soient trouvables et utilisables dans le cadre de l'enseignement actuel de l'OU et d'autres activités. La préservation est un sous-produit du projet plutôt qu'une fin en soi. Ce projet fournit un exemple important de combinaison de la préservation du contenu et de son utilisation, précieux pour l'institution, afin d'obtenir un budget et de produire un bénéfice. Le projet a été présenté lors de la journée d'information DPC « Préserver le son et la vidéo numériques ». Il a permis de numériser 1 200 bandes vidéo et films, et de placer les résultats dans un dépôt Fedora. En outre, 145 000 pages de documentation ont été numérisées, fournissant le cadre éducatif global autour des 1 200 contenus, leur donnant un contexte et améliorant leur capacité à être réutilisés. L'interface utilisateur permet une navigation granulaire et temporelle. Dans l'ensemble, ce projet est un exemple exceptionnel de bonnes pratiques.

Étude de cas n° 2 : projet d'enregistrements sonores des archives de la *British Library*

Il s'agit d'un projet de préservation et d'accès à l'éducation soutenu par le Jisc qui s'est déroulé (dans sa phase initiale) de 2004 à 2006. Une deuxième phase a permis d'ajouter des objets supplémentaires. Près de 50 000 enregistrements de discours, de musique et de sons de l'environnement humain et naturel ont été numérisés et mis en ligne. Le catalogue en ligne est ouvert à tous, et les établissements d'enseignement supérieur ou de formation continue britanniques agréés peuvent également écouter les enregistrements. Tout le monde peut écouter 2 000 de ces documents (ou n'importe lequel d'entre eux en se rendant dans la salle de lecture de la *British Library* à Londres). Les différences d'accès entre les établissements d'enseignement et le grand public reflètent la question globale des droits comme étant la seule contrainte qui subsiste sur le libre accès aux documents audiovisuels dans les institutions publiques.

Étude de cas n° 3 : projet PSRE de l'*Imperial War Museum*

L'*Imperial War Museum* possède l'une des plus importantes collections de films du Royaume-Uni. Il collectionne des films depuis sa fondation en 1919, en commençant par des images de la Grande Guerre qui ont conduit à la création de l'institution. Le fonds PSRE (*Public Sector Research Exploitation*) a accordé une subvention de près d'un million de livres sterling pour le catalogage, la numérisation et l'accès en ligne (au catalogue et aux images). Le projet s'est déroulé de 2006 à

2009 et présente un intérêt particulier dans la mesure où il vise spécifiquement l'exploitation commerciale d'une collection, et des modèles commerciaux durables autour de la numérisation et de l'accès au web. Le résultat est un site web (http://film.iwmcollections.org.uk/) où chacun peut consulter du contenu en basse qualité, mettre des documents, des photos et des images clés dans une galerie et remplir un panier d'achat pour ensuite acheter du contenu.

Étude de cas n° 4 : projet Newsfilm Online du British University Film and Video Council

Il s'agit d'un autre projet parrainé par le Jisc. Pendant quatre décennies jusqu'en 1960, les films d'actualités diffusés dans les salles de cinéma étaient le principal moyen pour le grand public de voir des images animées sur les événements récents. Le projet initial s'est déroulé de 2004 à 2008. Les résultats sont disponibles sur un site web qui, comme pour le projet BL *Archival Sound Recordings*, dispose de toutes les fonctionnalités pour les universités et les écoles supérieures enregistrées. Le grand public peut voir le catalogue complet et peut voir une seule vignette pour chaque article. Depuis la phase initiale du projet, le contenu a été enrichi par des nouvelles de ITN/Reuters couvrant les événements survenus des décennies après le déclin des films d'actualités. Les sujets des actualités sont courts : le projet initial fournissait 3 000 heures de contenu, mais cela représentait 60 000 sujets. En outre, comme pour le projet *Open University*, la documentation a également été mise en ligne pour le contexte et pour faciliter la recherche et l'extraction, soit 450 000 pages de scripts de bulletins d'information.

Étude de cas n° 5 : le *British Film Institute* (BFI) et le projet *Screen Heritage UK (*SHUK) des archives cinématographiques régionales

SHUK est un projet important (22,8 millions de livres sterling) et complexe (impliquant 12 services d'archives cinématographiques régionales en plus du BFI). Le projet a été compliqué par des changements dans la structure et le financement du BFI, ainsi que par un changement de gouvernement et une série d'autres problèmes. Néanmoins, le projet a produit des résultats importants :

- La préservation, et non la numérisation : construction d'une chambre forte de 6 millions de livres sterling pour la préservation des films ;
- La numérisation : numérisation des films et équipement de stockage numérique pour les archives cinématographiques régionales ;
- L'accès : catalogues en ligne du contenu des archives cinématographiques régionales, accessibles au grand public.

SHUK a été lancé le 5 septembre 2011 avec une production conjointe BBC et BFI, *The Reel History of Britain* (SHUK, 2011).

Conclusions

Le problème technologique de base pour les collections d'images animées et de sons réside dans la nécessité de numériser tous les contenus actuellement conservés sur des rayonnages. L'audio et la vidéo ont besoin de la numérisation pour leur survie même, en raison de l'obsolescence et de la dégradation des éléments physiques, qu'ils soient analogiques ou numériques. Les films peuvent être conservés (sauf s'ils se détériorent déjà) mais doivent être numérisés pour être accessibles.

La lecture pour une numérisation de préservation implique la nécessité d'une récupération optimale de la qualité originale, ce qui nécessite un équipement et une expérience professionnels. Le principal obstacle technique est que, pour de nombreux formats physiques, l'équipement nécessaire est en grande partie obsolète, ce qui signifie que les pièces et les réparations ainsi que les opérateurs qualifiés sont de plus en plus rares. La recommandation urgente est la suivante : n'attendez pas ! Les fonds audiovisuels doivent être documentés et faire partie d'un plan de préservation.

La situation du patrimoine sonore est claire. Les normes de numérisation, l'encodage, l'habillage et les métadonnées sont tous contenus et bien documentés dans l'IASA TC-04 (IASA, 2009). Le son non compressé dans le format BWF (*Broadcast Wave Format*) est largement utilisé et bien pris en charge. Il n'y a aucune raison de modifier le codage de base, bien que le format BWF puisse devenir obsolète à terme. Le seul problème important est l'incapacité de certaines applications audio standard à traiter correctement les métadonnées BWF intégrées (ARSC, 2011). Tous les services d'archives doivent être conscients du risque de perte des métadonnées intégrées. La situation de la vidéo est complexe, mais il existe une feuille de route PrestoSpace pour guider les choix en matière de numérisation des différents formats existants. Le PrestoCentre et le *Jisc Digital Media* donnent des conseils sur la préservation numérique des fichiers obtenus. Un des grands défis est d'établir un registre des applications qui fonctionnent correctement sur les métadonnées vidéo intégrées, où la diversité est énorme. Il n'existe pas de norme unique convenue pour les conteneurs, les métadonnées ou même le codage, et le passage de la définition standard à la haute définition apporte un nouvel ensemble d'applications, de conteneurs et de méthodes de codages.

Il existe une technologie émergente qui peut améliorer les transferts audio (capture du signal de polarisation et élimination conséquente de la variation temporelle) et vidéo (numérisation directe du signal RF de la tête de lecture), ce qui pourrait être utile dans les cas où la technologie actuelle échoue. Il est donc recommandé de ne pas attendre que cette technologie soit plus avancée et plus largement disponible. Si des problèmes de lecture ne peuvent être résolus, le format audio ou vidéo original doit être conservé afin que cette technologie avancée puisse être appliquée à l'avenir.

Le contrôle de la qualité des résultats de la numérisation reste un problème pour la vidéo. Il est nécessaire de combiner efficacement la technologie de traitement du signal avec le contrôle humain afin de produire une méthode vraiment efficace de contrôle de la qualité dans le cadre d'une approche industrielle de la préservation. Le contrôle qualité est tout aussi important pour la préservation numérique : toute modification ou migration due à l'obsolescence numérique doit être vérifiée pour préserver la qualité du signal. Là encore, une approche purement manuelle n'est pas adaptée (aux dizaines de millions d'heures de contenu audiovisuel dans les collections européennes), tandis que les substituts purement algorithmiques de type « recherche et écoute » n'ont jamais été complètement efficaces et restent un domaine où des recherches supplémentaires sont nécessaires.

Ressources

Wright, R., 2012. Preserving Moving Pictures and Sound DPC Technology Watch Report 12-01, mars 2012
http://dx.doi.org/10.7207/twr12-02
Ce rapport s'adresse à toute personne responsable de collections de contenus audiovisuels et ayant un intérêt dans la préservation de ces contenus. Les nouveaux contenus sont nativement numériques, l'audio et la vidéo analogique ont besoin de la numérisation pour survivre et les films ont besoin de la numérisation pour être accessibles. Par conséquent, la préservation numérique sera pertinente au fil du temps pour tous ces domaines. Le rapport se concentre sur la numérisation, l'encodage, les formats de fichiers et les enveloppes, l'utilisation de la compression, l'obsolescence et ce qu'il faut faire face aux problèmes particuliers de la préservation numérique du son et des images en mouvement (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
Communiqué de presse du BFI. 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
C'est le guide standard de la numérisation du son, et les sections sur les métadonnées et le stockage numérique sont utiles pour toutes les formes de médias numériques.

Casey, M. and Gordon, B., 2007. Best Practices for Audio Preservation. Bloomington, Indiana University Bloomington.
http://www.dlib.indiana.edu/projects/sounddirections/papersPresent/
Une autre ressource audio (qui comprend également une série de logiciels de numérisation) provient du projet *Sound Directions* des universités de Harvard et de l'Indiana : beaucoup de choses sont également pertinentes pour la numérisation de la vidéo. (160 pages)

Digital Preservation Coalition Journée d'information sur la préservation du son et de l'image numériques, avril 2011
https://www.dpconline.org/events/past-events/preserving-digital-sound-and-vision-a-briefing
Cette journée d'information de la DPC en avril 2011 a permis d'examiner et de débattre des derniers développements en matière de préservation du son et de l'image numériques. Sept présentations (dont l'Université ouverte) sont liées au programme et peuvent être téléchargées.

Comité technique de l'ARSC, 2011.*A Study of Embedded Metadata Support in Audio Recording Software: Summary of Findings and Conclusions*. *Association of Recorded Sound Collections*.
http://www.arsc-audio.org/pdf/ARSC_TC_MD_Study.pdf
Une étude sur la prise en charge des métadonnées intégrées dans et à travers une variété de logiciels d'enregistrement audio. Les conclusions soulèvent de sérieuses préoccupations, notamment pour les communautés d'archivage et de préservation qui dépendent des métadonnées intégrées pour l'interprétation et la gestion des fichiers numériques représentant le contenu préservé à l'avenir (21 pages).

AVPreserve
http://www.avpreserve.com/
Société de conseil en gestion des médias et de l'information basée aux États-Unis. Son site web propose une série de ressources pour la préservation de l'audiovisuel.

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
Le *British Film Institute* peut donner des conseils sur le film et aussi sur la vidéo – ils détiennent beaucoup de vidéo, et ont un conservateur pour la télévision. Sa mission est la collecte et la préservation des films et des émissions de télévision, ainsi que le conseil technique.

British Library Sound Archive
https://www.bl.uk/subjects/sound#
Des conseils techniques généraux sur la préservation du son sont disponibles auprès de la *British Library Sound Archive*. Elle a pour mission de collecter et de préserver toutes les formes d'audio, et de fournir des conseils techniques.

Film Archives UK
http://filmarchives.org.uk
Collecte et préservation du contenu audiovisuel général d'importance nationale au Royaume-Uni

Jisc Digital Media
https://www.jisc.ac.uk/
Conseils et formation sur les images fixes, les images animées et le son. Cela inclut leurs InfoKits pour les formats de fichiers numériques, le financement et la durabilité de la numérisation, et le Guide de numérisation de haut niveau pour les ressources audiovisuelles.

PrestoCentre
http://www.prestocentre.eu
Le site web fournit des informations, des ressources et des conseils dans le domaine de l'audiovisuel. L'accès a récemment été étendu de sorte que toutes les ressources sont désormais accessibles gratuitement à tous.

Maintenir une présentation vidéo cohérente
http://www.tate.org.uk/research/publications/sustaining-consistent-video-presentation
Ce document technique traite des approches permettant d'identifier et d'atténuer les risques liés au maintien de la présentation cohérente des fichiers vidéo numériques. Issu de deux projets de recherche multipartenaires - Pericles et Presto4U - le document a été commandé par *Tate Research* et est destiné à ceux qui sont activement engagés dans la préservation de la vidéo numérique.

–Jisc 2009 - Showreel d'enregistrements sonores d'archives
https://www.youtube.com/watch?v=KPy9ZqWEHog
Courte vidéo engageante sur le projet d'enregistrements sonores d'archives de la *British Library* publiée le 22 juin 2009. (6 minutes 11 secondes).

Pour aller plus loin

Podcasts dans les archives : Archivage des contenus de podcasts à l'Université du Michigan
http://files.archivists.org/pubs/CampusCaseStudies/CASE12.pdf
Dans cette étude de cas des services d'archives des universités de la *Society of American Archivists* (Campus Case Studies), Alexis. A. Antracoli, de l'Université du Michigan, examine les défis liés au développement des bonnes pratiques (best practices) et des processus de travail pour l'archivage et la préservation des contenus de podcast. L'un des principaux enjeux consistait à établir des normes de pratique pour la capture, le stockage et l'accès, en particulier la génération et le stockage de métadonnées descriptives, techniques et de préservation appropriées. Un autre défi a consisté à développer l'infrastructure technologique nécessaire pour soutenir un système conforme à l'*Open Archives Information System* (OAIS). 2010 (14 pages).

Références

ARSC Technical Committee, 2011.*Study of Embedded Metadata Support in Audio Recording Software*. Association of Recorded Sound Collections. Disponible à l'adresse suivante : 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. Disponible à l'adresse suivante : http://www.iasa-web.org/audio-preservation-tc04

SHUK, 2011.*Screen Heritage UK Marks new Era for Britain's Film Archives.* Disponible à l'adresse suivante : 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. Disponible à l'adresse suivante : http://dx.doi.org/10.7207/twr12-02

Archivage web

Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Aperçu

Cette étude de cas, résumé du rapport de veille technologique de la DPC au sujet de l'archivage web, donne un bref aperçu de niveau débutant à intermédiaire. Trois « mini études de cas » sont incluses pour illustrer les différents contextes opérationnels, les moteurs et les solutions qui peuvent être mis en œuvre. Le rapport lui-même constitue une étude approfondie qui aborde un plus large éventail de problématiques et de pratiques, avec une bibliographie et des conseils supplémentaires (Pennock, 2013). Il est recommandé aux lecteurs qui sont à la recherche d'un rapport de niveau plus avancé sur le sujet et sa mise en œuvre.

Introduction

Le web est une ressource d'information unique et de grande envergure, utilisée à l'échelle mondiale. Une grande partie de son contenu aura probablement une valeur non seulement pour la génération actuelle, mais aussi pour les générations futures. Pourtant, l'héritage durable du web est en danger, menacé en partie par la vitesse même à laquelle il est devenu un succès. Les contenus se perdent à un rythme alarmant, ce qui met en péril non seulement notre mémoire culturelle numérique, mais aussi la responsabilité des organisations. Conscientes de ce fait, un certain nombre d'institutions universitaires, d'organisations à but non lucratif et d'entreprises privées ont exploré les questions en jeu et ont dirigé ou contribué à l'élaboration de solutions techniques pour l'archivage web.

Services et solutions

Les besoins des entreprises et les ressources disponibles sont fondamentaux lors de la sélection d'outils et/ou de services d'archivage web appropriés. D'autres questions connexes doivent également être prises en compte : les organisations qui envisagent l'archivage web pour répondre aux exigences réglementaires doivent, par exemple, prendre en compte des questions connexes telles que l'authenticité et l'intégrité, la politique d'archivage de documents et l'assurance qualité. Toutes les organisations devront se pencher sur la question de la sélection (c'est-à-dire quels sites web archiver), une tâche apparemment simple qui est compliquée à cause des interrelations complexes que partagent la plupart des sites web et qui rendent difficile la fixation de limites. Parmi les autres questions, citons la gestion des logiciels malveillants, la réduction au minimum de

la duplication des ressources, la cohérence temporelle des sites et la préservation ou la durabilité à long terme des ressources. La collaboration internationale s'avère être un facteur déterminant en matière de développement de solutions évolutives pour soutenir la préservation à long terme et garantir que les collections restent accessibles de manière fiable pour les générations futures.

Le processus d'archivage web n'est pas une action ponctuelle. Une suite d'applications est généralement déployée pour soutenir les différentes étapes du processus, bien qu'elles puissent être intégrées dans un processus de travail unique de bout en bout. Une grande partie des logiciels est disponible en Open Source, ce qui permet aux institutions d'accéder librement au code source pour l'utiliser et/ou le modifier gratuitement.

Systèmes intégrés pour l'archivage web

Un petit nombre de systèmes intégrés sont disponibles pour ceux qui disposent de ressources techniques suffisantes pour installer, maintenir et administrer un système en interne. Ces systèmes offrent généralement des fonctionnalités d'archivage web intégrées sur la majeure partie du cycle de vie, de la sélection et de la gestion des autorisations à la navigation, l'assurance qualité et l'accès. Trois systèmes sont présentés ici.

PANDAS
PANDAS (*PANDORA Digital Archiving System*) a été l'un des premiers systèmes d'archivage web intégrés disponibles. D'abord mis en œuvre par la Bibliothèque nationale d'Australie (NLA) en 2001, PANDAS est une application web écrite en Java et Perl qui fournit une interface conviviale pour gérer les processus de l'archivage web. Elle prend en charge la sélection, les autorisations, la programmation, l'exploration, l'assurance qualité, la préservation et l'accès. PANDAS n'est pas un logiciel Open Source, bien qu'il ait été utilisé par d'autres institutions (notamment le *UK Web Archiving Consortium* de 2004 à 2008). Il est utilisé par la NLA pour l'archivage web sélectif, tandis qu'Internet Archive prend en charge les captures annuelles du domaine australien.
Web Curator Tool (WCT)
Le *Web Curator Tool* est un outil de processus de travail Open Source pour la gestion du processus d'archivage web sélectif, développé en collaboration par la Bibliothèque nationale de Nouvelle-Zélande et la *British Library* avec *Oakleigh Consulting*. Il prend en charge la sélection, les autorisations, la description, la collecte et l'assurance qualité, avec une interface d'accès séparée. WCT est écrit en Java au sein d'une architecture flexible et peut être téléchargé sur SourceForge sous une licence publique Apache. Le site web de WCT est au centre de la communauté des développeurs et il existe des listes de diffusion actives pour les utilisateurs et les développeurs. La nature hautement modulaire du système minimise les dépendances.
NetarchiveSuite
NetarchiveSuite est une application d'archivage web écrite en Java pour gérer l'archivage web sélectif et à large domaine, développée à l'origine en 2004 par les deux bibliothèques de dépôt légal au Danemark (*Det Kongelige Bibliotek* et *Statsbiblioteket*). Elle est devenue Open Source en 2007 et a reçu des contributions de développement supplémentaires de la Bibliothèque nationale de France et de l'*Österreichische Nationalbibliothek* depuis 2008. Elle est disponible gratuitement sous licence GNU *Lesser General Public License* (LGPL). La nature hautement modulaire du système permet des solutions de mise en œuvre flexibles.

Services tiers et services commerciaux

Les services tiers d'archivage web commerciaux sont de plus en plus utilisés par les organisations qui préfèrent ne pas établir et maintenir leur propre infrastructure technique d'archivage web. Les raisons de ce choix peuvent varier considérablement. Souvent, il ne s'agit pas simplement de l'ampleur de l'opération ou de la complexité perçue, mais des besoins et des objectifs propres à l'entreprise. De nombreuses organisations ne souhaitent pas investir dans des compétences ou des capitaux qui ne sont pas essentiels à leur activité. D'autres peuvent utiliser un tel service pour éviter l'investissement en capital. En outre, les organisations déplacent de plus en plus leurs opérations et activités informatiques dans le *cloud*, ou utilisent un fournisseur de SAAS (*Software as a Service*). L'archivage web ne fait pas exception à la règle. D'un point de vue juridique et de conformité, les services tiers sont parfois préférés car ils peuvent fournir non seulement la technologie, mais aussi les compétences et le soutien nécessaires pour répondre aux besoins des entreprises. Cette section présente certains des services tiers actuellement disponibles, mais il s'agit bien sûr d'une liste non exhaustive, et son inclusion ici ne doit pas être considérée comme une recommandation.

Archive-It
Archive-It est un service d'archivage web par abonnement fourni par Internet Archive. Les clients utilisent le service pour établir des collections spécifiques, par exemple sur les Jeux Olympiques de Londres de 2012, sur les sites web gouvernementaux, sur les droits de l'homme ou sur des bibliographies de cours. Une interface utilisateur dédiée permet aux clients de sélectionner et de gérer les URL de départ (*seeds*), de définir la profondeur, l'étendue et la fréquence de l'exploration, de surveiller la progression de l'exploration et d'effectuer l'assurance qualité, d'ajouter des métadonnées et de créer des pages de destination pour leurs collections. Les collections sont rendues publiques par défaut via le site web d'Archive-It, les collections privées nécessitant une disposition spéciale. L'interface d'accès prend en charge la recherche par URL et par texte intégral. Plus de 200 partenaires utilisent le service, la plupart issus des secteurs universitaires ou du patrimoine culturel. Le coût du service dépend des exigences de l'institution de collecte.
Archivethe.Net
Archivethe.Net est un service d'archivage de sites web fourni par la *Fondation Internet Memory* (IMF). Il permet aux clients de gérer l'ensemble du processus de travail via une interface web selon trois modules principaux : administration (gestion des utilisateurs), collecte (gestion des URL de départ [*seeds*] et de l'exploration) et rapport (rapports et mesures à différents niveaux). La plate-forme est disponible en anglais et en français. Outre la recherche en texte intégral et la collecte de contenu multimédia, elle prend également en charge un service de redirection automatisé pour les sites en ligne. Des outils automatisés d'assurance qualité sont en cours de développement, mais l'IMF peut aussi fournir des services de contrôle qualité manuelle, ainsi que la gestion directe de la collecte pour les institutions qui ne souhaitent pas utiliser l'outil en ligne. Les coûts dépendent des exigences de l'institution de collecte. Les collections peuvent être rendues privées ou rester librement accessibles, auquel cas elles peuvent porter des marques selon les exigences des institutions de collecte et apparaître dans la collection de l'IMF. Dans ce cas, les frais d'hébergement sont pris en charge par l'IMF.
The University of California's *Curation Centre* (UC3)
Dans le cadre de la *California Digital Library*, l'UC3 fournit un service d'archivage web entièrement hébergé pour des collections d'archives web sélectives. Les départements et organisations de l'Université de Californie ne sont facturés que pour le stockage. Des frais sont

perçus auprès des autres groupes et consortiums, comprenant une redevance annuelle de service plus les coûts de stockage. Les collections peuvent être mises à la disposition du public ou conservées à titre privé. À ce jour, une vingtaine d'organisations partenaires ont mis des collections à disposition. La recherche en texte intégral est assurée et la présentation des collections peut porter des marques selon les besoins des institutions de collecte.
Entreprises privées
Les entreprises privées proposent des services d'archivage de sites web particulièrement adaptés aux besoins des entreprises. Hanzo Archives, par exemple, fournit un service d'archivage de sites web commerciaux pour répondre aux besoins des entreprises commerciales en matière de conformité réglementaire, d'e-discovery et de gestion des documents. Hanzo Archives met l'accent sur sa capacité à collecter des sites et des contenus interactifs (*rich media*) qui peuvent être difficiles à trouver pour un robot d'exploration standard, notamment des contenus dynamiques de Sharepoint et des wikis d'internautes privés, ainsi que des canaux de réseaux sociaux publics et privés (pour plus de détails sur les possibilités offertes par le service Hanzo Archives, voir l'étude de cas Coca-Cola). De même, Reed Archives fournit un service commercial d'archivage web pour la conformité réglementaire des organisations, la protection des litiges, l'e-discovery et la gestion des documents. Ce service comprend un ensemble d'outils « d'archivage à la demande » à utiliser lors de la navigation sur le web. Dans chaque cas, le coût du service est adapté aux besoins précis du client. D'autres sociétés et services sont également disponibles et les lecteurs sont encouragés à rechercher en ligne d'autres options si un tel service présente un intérêt.

Étude de cas n°1 : *UK Web Archive*

Les *UK Web Archive* (UKWA) ont été créées en 2004 par le *UK Web Archiving Consortium*. Il s'agissait à l'origine d'un partenariat à six, mené par la *British Library* en collaboration avec la *Wellcome Library*, Jisc, la Bibliothèque nationale du Pays de Galles, la Bibliothèque nationale d'Ecosse et les Archives nationales du Royaume-Uni.

Les partenaires de l'UKWA sélectionnent et désignent les sites web en utilisant les fonctionnalités du système d'archivage web hébergé sur l'infrastructure des *UK Web Archive* maintenue par la *British Library*. La *British Library* travaille en étroite collaboration avec un certain nombre d'autres institutions et de personnes pour sélectionner et nommer des sites web dignes d'intérêt. Les sites web archivés de manière sélective sont revus à intervalles réguliers afin de garder la trace des changements intervenus au fil du temps.

L'infrastructure technique sur laquelle repose les *UK Web Archive* est gérée par la *British Library*. Les *UK Web Archive* ont été initialement créées avec le logiciel PANDAS fourni par la Bibliothèque nationale d'Australie, hébergé par une agence externe, mais en 2008, les *UK Web Archive* ont été déplacées en interne et ont migré vers le système *Web Curator Tool* (WCT).

Une version personnalisée de l'interface Wayback développée par l'Internet Archive est utilisée comme interface WCT et permet d'accéder à tous les sites web archivés accessibles au public. La recherche en texte intégral est activée en plus des recherches standard de titres et d'URL et d'un schéma de classification par sujet. L'équipe d'archivage web de la bibliothèque a récemment mis en ligne un certain nombre d'outils de visualisation pour aider les chercheurs à comprendre et à trouver le contenu de la collection.

Des collections spéciales ont été créées sur un large éventail de sujets. Nombre d'entre elles sont thématiques, par exemple la santé mentale et les collections de l'Église libre d'Ecosse. D'autres documentent la réponse en ligne à un événement notable de l'histoire récente, comme les élections générales au Royaume-Uni, le jubilé de diamant de la reine Elizabeth II et les Jeux olympiques de Londres 2012.

De nombreux autres sites, qui ne sont pas associés à une collection spéciale donnée, ont été archivés sur recommandation de spécialistes du sujet ou de membres du public. Ces sites ne sont souvent plus disponibles sur le web, comme par exemple le site du député britannique Robin Cook ou le projet d'art public *One & Other* d'Antony Gormley, acquis auprès de Sky Arts.

Étude de cas n° 2 : Internet Memory Fundation

La *Internet Memory Fundation* (IMF) a été créée en 2004 en tant qu'organisation à but non lucratif pour soutenir les initiatives d'archivage web et développer le soutien à la préservation du web en Europe. Connue à l'origine sous le nom de *European Archive Foundation*, elle a changé de nom en 2010. L'IMF fournit à ses clients une solution d'archivage web externalisée à part entière pour gérer le processus de travail d'archivage web sans qu'ils aient à faire face à des problèmes opérationnels.

L'IMF collabore étroitement avec l'*Internet Memory Research* (IMR) pour exploiter une partie de ses processus de travail techniques pour l'archivage web. L'IMR a été créée en 2011 en tant que société dérivée de L'IMF. L'IMF et l'IMR sont impliqués dans des projets de recherche qui soutiennent la croissance et l'utilisation des archives web.

L'IMR fournit un service d'archivage web personnalisable, Archivethe.Net (AtN). AtN est une plateforme d'archivage web dotée d'une interface web qui aide les institutions à débuter facilement et rapidement la collecte des sites web incluant du contenu dynamique et interactif (*rich media*). Elle peut être adaptée aux besoins des clients, et les institutions conservent le contrôle total de leur politique de collecte (possibilité de sélectionner les sites, de spécifier la profondeur, la fréquence de collecte, etc.) Des services de contrôle de la qualité peuvent être fournis sur demande. La plupart sont effectués manuellement afin de répondre aux exigences de qualité élevées des institutions. L'IM dispose d'une équipe dédiée à l'assurance qualité, composée d'évaluateurs d'assurance qualité. L'IM a développé une méthodologie de comparaison visuelle basée sur les outils utilisés pour l'exploration et l'accès aux données, bien qu'elle travaille également à l'amélioration des outils et des méthodes afin de fournir une meilleure qualité initiale d'exploration.

Parmi les institutions partenaires qui possèdent des collections librement accessibles et pour lesquelles l'IM fournit un service d'archivage web figurent les Archives nationales du Royaume-Uni et le Parlement britannique.

L'accès aux collections accessibles au public est assuré par le site web de l'IM. L'IM offre un service de recherche en texte intégral pour la plupart de ses collections en ligne, en plus de la recherche par URL. Les résultats des recherches en texte intégral peuvent être intégrés sur un site web tiers et les collections peuvent porter des marques si nécessaire.

Conformément à l'architecture du service de continuité du web (*Web Continuity Service*) des Archives nationales du Royaume-Uni (*The National Archives*, 2010), l'IM a mis en place un service de redirection automatique pour intégrer les archives web à l'expérience des utilisateurs du web. En naviguant sur le web, les utilisateurs sont automatiquement redirigés vers les archives web si la ressource demandée n'est plus disponible en ligne. Dans les archives web, l'utilisateur est dirigé vers l'instance la plus récente de la ressource demandée. Une fois que la ressource est accessible, tout lien sur la page renvoie l'utilisateur à la version en ligne du site. Ce service permet d'augmenter la durée de vie d'un lien, améliorant l'expérience des utilisateurs, la visibilité et le classement en ligne, et réduisant les taux de rebond.

Les collections d'archivage web sont disponibles pour la navigation publique à partir du site web de l'IM, une combinaison de collections de domaines et de collections sélectives provenant de ses propres institutions et d'institutions partenaires.

Étude de cas n° 3 : les archives web de Coca-Cola

Les archives web de Coca-Cola ont été créées pour saisir et préserver les sites web et les médias sociaux de l'entreprise Coca-Cola. Elles font partie des archives de Coca-Cola, qui contiennent des millions d'objets physiques et numériques, allant des papiers et des photographies aux publicités, aux bouteilles et aux articles promotionnels. La présence en ligne de Coca-Cola est vaste et comprend non seulement plusieurs sites web nationaux de Coca-Cola, mais aussi, par exemple, la page Facebook et le flux Twitter de Coca-Cola, ainsi que d'autres marques appartenant à Coca-Cola (500 au total). Le premier site web de Coca-Cola a été publié en 1995.

Depuis 2009, Coca-Cola collabore avec *Hanzo Archives* et utilise désormais leur service commercial d'archivage de sites web. Outre les avantages patrimoniaux des archives web, le service fournit également un soutien en cas de litiges lorsqu'une partie ou la totalité du site web peut être invoquée comme preuve devant un tribunal. Il fournit également une conformité réglementaire pour les applications de gestion des documents.

Les archives web de Coca-Cola sont des archives web thématiques spéciales qui contiennent tous les sites commerciaux de Coca-Cola et d'autres sites spécialement sélectionnés associés à Coca-Cola. Elles se veulent aussi complètes que possible, l'intégrité et la fonctionnalité des sites capturés étant de première importance. Cela inclut les médias sociaux et les vidéos, qu'ils soient diffusés en direct ou intégrés (y compris Flash). Les objets sont préservés dans leur forme originale dans la mesure du possible, un principe fondamental pour toutes les archives de Coca-Cola.

Les robots d'exploration de *Hanzo Archives* interviennent tous les trimestres et sont complétés par des explorations occasionnelles basées sur des événements, comme le 125e anniversaire de Coca-Cola, célébré en 2011. La solution d'archivage web de Hanzo est une application personnalisée. Le contenu web est collecté dans son format natif par le robot d'exploration du web de *Hanzo Archives*, qui est déployé à l'échelle nécessaire pour la tâche en question.
L'assurance qualité est effectuée grâce à un contrôle systématique par échantillonnage en deux bonds du contenu de l'exploration qui force l'utilisation des options de navigation de niveau supérieur et se concentre sur la forme technique du site.

Les archives sont actuellement accessibles uniquement aux employés de Coca-Cola, sur un nombre limité de machines. L'accès à distance est assuré par Hanzo, qui utilise sa propre interface d'accès.

L'accès proxy garantit que tout le contenu est diffusé directement à partir des archives et qu'il n'y a pas de « fuite du site en direct ». Les archives pourraient être rendues accessibles au public à l'avenir dans le musée World of Coca-Cola, à Atlanta (États-Unis, Géorgie).

La collection d'archives web de Coca-Cola contient plus de six millions de pages web et plus de 2 To de données. Avant sa collaboration avec Hanzo, les premières tentatives d'archivage ont abouti à des captures incomplètes, de sorte que les premiers sites ne sont pas aussi complets que la société le souhaiterait. La collection contient également des informations sur de nombreux événements nationaux et internationaux pour lesquels Coca-Cola était sponsor, notamment les Jeux olympiques de Londres 2012 et le Jubilé de diamant de la Reine Elizabeth II.

Conclusions

La technologie d'archivage web a considérablement évolué au cours de la dernière décennie, tout comme notre compréhension des enjeux. Par conséquent, nous disposons d'un large éventail d'outils et de services qui nous permettent d'archiver et de préserver certains aspects de notre mémoire culturelle en ligne, ainsi que de nous conformer aux exigences réglementaires en matière de saisie et de préservation des documents en ligne. Le travail est en cours, car tant que l'Internet continuera à évoluer, la technologie d'archivage web devra évoluer pour suivre le rythme.

Parallèlement aux développements techniques, les connaissances et l'expérience acquises grâce au déploiement et à l'utilisation pratiques des outils d'archivage web ont permis de mieux comprendre quelles étaient les meilleures pratiques en matière d'archivage web, les stratégies opérationnelles pour intégrer l'archivage web dans un contexte organisationnel, les besoins et les avantages commerciaux, les cas d'utilisation et les différentes options en matière de ressources. Les organisations qui souhaitent se lancer dans une initiative d'archivage web doivent être très claires quant à leurs besoins commerciaux. Les besoins de l'entreprise doivent être le moteur fondamental de toute initiative d'archivage web et influenceront considérablement les détails de la stratégie d'archivage web et de la politique de sélection qui en résulteront. L'émergence de technologies et de services commerciaux est un signe de la maturité de l'archivage web en tant que besoin commercial, ainsi que discipline.

Ressources

Pennock, M., 2013. Web-Archiving, DPC Technology Watch Report 13-01 Mars 2013
http://dx.doi.org/10.7207/twr13-01
Ce rapport est destiné à ceux qui s'intéressent à la mise en place d'une archive web ou qui en ont la responsabilité. Il présente et examine les principales questions auxquelles sont confrontées les organisations engagées dans des initiatives d'archivage de sites web, qu'elles sous-traitent à un prestataire de services tiers ou qu'elles gèrent le processus en interne, et donne un aperçu détaillé des principales applications et outils logiciels actuellement disponibles.

ISO, 2012, ISO 28500:2009 Information and Documentation – the WARC file format (disponible en anglais uniquement)
http://www.iso.org/iso/catalogue_detail.htm?csnumber=44717

Le format WARC (Web ARChive) est un format de conteneur pour les sites web archivés, également connu sous le nom d'ISO 28500:2009. Il s'agit d'une révision du format de fichier ARC de l'Internet Archive utilisé pour stocker les explorations du web collectées sur le World Wide Web.

ISO, 2013 ISO/TR 14873:2013 Information and Documentation – Statistics and quality issues for web archiving (disponible en anglais uniquement)
http://www.iso.org/iso/catalogue_detail.htm?csnumber=55211
Ce rapport technique définit des statistiques, des termes et des critères de qualité pour l'archivage web. Il examine les besoins et les pratiques d'un large éventail d'organisations telles que les bibliothèques, les archives, les musées, les centres de recherche et les fondations patrimoniales.

Meyer E 2010 (a), Researcher Engagement with Web Archives: State of the Art Report, Jisc
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1714997
Ce rapport résume l'état de l'art de l'archivage web par rapport aux chercheurs et aux besoins de recherche en se concentrant principalement sur les chercheurs individuels et les 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. Disponible aux adresses suivantes : http://ssrn.com/abstract=2329161, http://dx.doi.org/10.2139/ssrn.2329161,
http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2329161 ou
http://dx.doi.org/10.2139/ssrn.2329161
Cet article de l'équipe du projet Perma documente un grave problème de « pourriture des références » : plus de 70 % des URL de la *Harvard Law Review* et d'autres revues, et 50 % des URL des avis de la Cour suprême des États-Unis, ne renvoient pas aux informations initialement citées. Elle propose une solution aux auteurs et aux éditeurs de nouvelles études qui implique que les bibliothèques entreprennent la préservation à long terme des contenus des liens.

Scholarly Context Not Found: One in Five Articles Suffers from Reference Rot
http://dx.doi.org/10.1371/journal.pone.0115253
Cette étude à grande échelle a examiné environ 600 000 liens extraits de plus de 3 millions d'articles scientifiques publiés entre 1997 et 2012. Elle inclut de liens vers des ressources dites « web-at-large », c'est-à-dire des liens autres que des liens vers des articles scientifiques. Elle a révélé qu'un article scientifique, technique ou médical sur cinq souffrait de pourriture de références, ce qui signifie qu'il est impossible de revoir le contexte web qui les entoure quelque temps après leur publication. Si l'on considère uniquement les articles scientifiques, techniques ou médicaux qui contiennent des références à des ressources web, cette fraction passe à sept sur dix.

The national archives du Royaume-Uni, 2010. *Government web archiving* : *Redirection Technical Guidance for Government Departments*, version 4.2, Archives nationales du Royaume-Uni
http://www.nationalarchives.gov.uk/documents/information-management/redirection-technical-guidance-for-departments-v4.2-web-version.pdf
Ce guide décrit un service innovant qui offre une fonctionnalité de réécriture et de redirection d'URL pour les pages web du gouvernement britannique en mettant en place une redirection vers les archives web du gouvernement britannique lorsqu'une URL demandée n'existe plus sur un site web ministériel.

MEMENTO et le Time Travel Service
http://timetravel.mementoweb.org/
Memento est un outil qui permet aux utilisateurs de voir une version d'une ressource web telle qu'elle existait à un certain moment dans le passé. Il est maintenant utilisé dans plusieurs archives web. Le Time Travel Service basé sur Memento vérifie une série de serveurs comprenant de nombreuses archives web et essaie de trouver une page web telle qu'elle existait à l'époque de votre choix.

Archive-It
http://www.archive-it.org/

Hanzo Archives
http://www.hanzoarchives.com/

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

UC3 Web Archiving Service
https://cdlib.org/services/uc3/about/

Web Curator Tool
https://webcuratortool.org/

International Internet Preservation Consortium
http://www.netpreserve.org
L'IIPC est une organisation dont les membres se consacrent à l'amélioration des outils, des normes et des meilleures pratiques de l'archivage web tout en promouvant la collaboration internationale ainsi qu'un large accès et une utilisation des archives web pour la recherche et le patrimoine culturel. Le site web contient de nombreuses ressources précieuses, notamment d'excellentes vidéos courtes comme l'exemple ci-dessous.

Why Archive the Web?
https://www.youtube.com/watch?v=pU32rjTaM

Une courte vidéo publiée le 18 octobre 2012 présentant les défis de l'archivage web et de l'IIPC. (2 minutes 53 secondes).

What is a Web Archive?

https://youtu.be/ubDHY-ynWi0

Cette courte vidéo explique l'archivage web et pourquoi il est important que les bibliothèques de dépôt légal du Royaume-Uni le soutiennent. Elle a été produite dans le cadre du projet *Big UK Domain Data for the Arts and Humanities*, financé par le Conseil de la recherche sur les arts et les sciences humaines (2 min 31 s).

What do the UK Web Archive collect?

https://youtu.be/1QLMPIRwJEo

Cette vidéo destinée aux utilisateurs explique ce qu'ils peuvent s'attendre à trouver et où ils peuvent se rendre pour accéder aux trois collections que détiennent les *UK Web Archive*. Elle a été produite dans le cadre du projet *Big UK Domain Data for the Arts and Humanities*, financé par le Conseil de la recherche sur les arts et les sciences humaines (2 minutes 55 secondes).

Autres études de cas

Études de cas sur le contenu du site web de la NDSA

La *National Digital Stewardship Alliance* (NDSA) des États-Unis examine la valeur, les possibilités et les obstacles de la préservation sélective des types de contenus web spécifiques suivants :

Forums sur les sciences, la médecine, les mathématiques et la technologie

http://www.digitalpreservation.gov/ndsa/working_groups/documents/ScienceForums_CaseStudy_public_v2.pdf

décembre 2013 (3 pages).

Blogs sur les sciences, la médecine, les mathématiques et la technologie

http://www.digitalpreservation.gov/ndsa/working_groups/documents/ScienceBlogs_CaseStudy_public_v2.pdf

décembre 2013 (3 pages).

Born-Digital Community and Hyperlocal News

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

Février 2013 (3 pages).

Citizen Journalism

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

Février 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*, Université du Michigan, détaille les stratégies et les procédures suivies par le programme *University Archives and Records Program* (UARP) pour développer sa collection de sites web archivés, et comment il a lancé un projet de préservation des

sites web à grande échelle dans le cadre d'un effort plus large visant à capturer et à maintenir de manière proactive certains documents électroniques de l'Université. 2011 (29 pages).

Références

Pennock, M., 2013.Web-Archiving, *DPC Technology Watch Report* 13-01 March 2013.Disponible à l'adresse suivante : http://dx.doi.org/10.7207/twr13-01

The national archives du Royaume-Uni, 2010. *Government Web Archive: Redirection Technical Guidance for Government Departments*, version 4.2, The National Archives (UK). Disponible à l'adresse suivante : http://www.nationalarchives.gov.uk/documents/information-management/redirection-technical-guidance-for-departments-v4.2-web-version.pdf

Glossaire

Illustration par Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Introduction

Les acronymes et les initiales sont une caractéristique de toute discipline spécialisée. Dans une discipline émergente, telle que la préservation numérique, une autre difficulté majeure est l'absence d'une taxonomie précise et définitive des termes. Différentes communautés utilisent les mêmes termes de différentes manières, ce qui peut faire obstacle à une communication efficace. Les définitions et les acronymes suivants sont ceux utilisés dans le manuel, les rapports de veille technologique et le site web de la DPC. Ils sont destinés à faciliter son utilisation en tant qu'outil pratique.

ABCDEFGHIJKLMNOP Q R STU V WX Y Z

A

Accès Comme défini dans le manuel, l'accès est à comprendre comme l'utilisation continue et permanente d'une ressource numérique, en conservant toutes les qualités d'authenticité, de précision et de fonctionnalité jugées essentielles en relation avec les finalités de création ou d'acquisition de l'objet numérique.

Accès après désabonnement *Post-cancellation Access,* voir Accès perpétuel.

Accès perpétuel *Continuing Access*. L'accès perpétuel fait référence au droit d'un abonné à une publication électronique et de ses utilisateurs d'avoir un accès permanent aux documents électroniques qui ont déjà été loués et payés par l'abonné auprès d'un éditeur. C'est un terme utilisé, avec ses synonymes d'accès permanent et d'accès après désabonnement, dans l'industrie de l'information pour décrire la capacité à conserver l'accès aux matériaux électroniques par l'abonné/le licencié après que l'accord de licence contractuel avec l'éditeur pour ces matériaux a pris fin, quelle que soit la raison de la cessation. Elle peut également couvrir, le cas échéant, les dispositions de préservation numérique nécessaires pour garantir certains éléments de l'accès perpétuel.

ADS *Archaeology Data Service* (Service de données archéologiques). Un service basé au Royaume-Uni et actif dans la préservation numérique. http://ads.ahds.ac.uk

AIP *Archival Information Package* (Paquet d'informations archivé). Un ensemble d'informations, comprenant les informations de contenu et les informations de description de la préservation (PDI) associées, qui sont préservées dans un système OAIS (terme de l'OAIS).

AMIA *Association of Moving Image Archives*. Une organisation active dans le domaine de l'archivage des images animées. http://www.amianet.org

ARC Format de conteneur pour les sites web conçu par l'Internet Archive, remplacé par le WARC (ISO 28500:2009).

Archivage numérique *Digital Archiving*. Ce terme est utilisé de manière très différente en fonction des secteurs. Les communautés des bibliothèques et des archives l'utilisent souvent de manière interchangeable avec la préservation numérique. Les professionnels de l'informatique ont tendance à utiliser l'archivage numérique pour désigner le processus de sauvegarde et de maintenance continue, par opposition aux stratégies de préservation numérique à long terme. C'est cette dernière définition plus riche, telle que définie dans le cadre de la préservation numérique, qui a été utilisée tout au long de ce manuel.

ASCII *American Standard Code for Information Interchange*. Norme pour le texte électronique. https://en.wikipedia.org/wiki/ASCII (en) ou https://fr.wikipedia.org/wiki/American_Standard_Code_for_Information_Interchange (fr).

Atelier et tutoriel sur la gestion de la préservation numérique. *Digital Preservation Management Workshop and Tutorial***.** Un atelier de formation intensive et un tutoriel en ligne développés et maintenus par la *Cornell University Library*, 2003-2006 ; étendus et maintenus par l'ICPSR, 2007-2012 ; et maintenant étendus et maintenus par les bibliothèques du MIT, 2012-en cours. http://dpworkshop.org/

Authentification Mécanisme qui tente d'établir l'authenticité des objets numériques à un moment donné. Par exemple, les signatures numériques.

Authenticité L'objet numérique est ce qu'il prétend être. Dans le cas des documents électroniques, elle fait référence à la fiabilité du document électronique en tant que document. Dans le cas d'objets « nés numériques » et numérisés, elle fait référence au fait que ce qui est cité est identique à ce qu'il était lors de sa création initiale, à moins que les métadonnées qui l'accompagnent n'indiquent des changements. La confiance dans l'authenticité des objets numériques au fil du temps est particulièrement cruciale en raison de la facilité avec laquelle des modifications peuvent être apportées.

B

Bac à sable *Sandbox Containment*. Environnement informatique sécurisé pour l'exécution de codes nouveaux, non testés ou expérimentaux ou de modifications de codes, y compris de codes potentiellement malveillants. L'environnement est autonome, ses ressources sont étroitement contrôlées et il est virtuel par essence.

Bit Un bit est l'unité de base de l'information en informatique. Il ne peut avoir qu'une seule des deux valeurs généralement représentées par un 0 ou un 1. Les deux valeurs peuvent être interprétées comme n'importe quel attribut à deux valeurs (oui/non, on/off, etc.).

Bit Preservation, voir Préservation du train de bits

Bloqueurs d'écriture *Writeblockers*. Outils qui empêchent un système informatique d'inspection d'écrire ou de modifier une collection ou un disque dur unitaire ou tout autre objet médiatique numérique. Les bloqueurs d'écriture matériels sont généralement considérés comme plus fiables que les bloqueurs d'écriture logiciels.

Born-Digital, voir Nativement numérique.

BWF *Broadcast WAV format*. La norme de l'Union européenne de radiodiffusion pour un fichier WAV, avec des métadonnées supplémentaires. http://www.digitalpreservation.gov/formats/fdd/fdd000003.shtml

Byte (B), voir Octet

C

CCSDS *Consultative Committee for Space Data Systems*, l'organisme responsable du modèle de référence OAIS. http://public.ccsds.org/default.aspx

Chaîne de confiance *Chain of Custody.* Concept clé en criminalistique selon lequel la garde et la provenance du matériel, des supports et des fichiers numériques sont sauvegardées grâce, par exemple, à la désignation de gardiens des preuves. L'objectif du sachet de preuves numériques (*Digital Evidence Bag*) est de conserver numériquement, avec les objets numériques probants, des métadonnées de provenance qui peuvent être mises à jour si nécessaire : un concept familier aux professionnels de la préservation numérique.

Checksum, voir Empreinte

CLIR *Council on Library and Information Resources*. Organisation basée aux États-Unis et active dans le domaine de la préservation numérique. http://www.clir.org

CNI *Coalition for Networked Information*. Organisation basée aux États-Unis et active dans la préservation numérique. http://www.cni.org

Communauté d'utilisateurs cible *Designated Community***.** Groupe identifié d'utilisateurs potentiels, susceptibles de comprendre un ensemble donné d'informations. La communauté d'utilisateurs cible peut être constituée de plusieurs communautés d'utilisateurs. La communauté d'utilisateurs cible est définie par l'Archive et sa définition peut évoluer au cours du temps (terme OAIS).

Compression avec pertes *Lossy Compression*. Mécanisme de réduction de la taille des fichiers qui supprime généralement une partie des données.

Compression sans perte *Lossless Compression*. Mécanisme de réduction de la taille des fichiers qui permet de conserver toutes les données originales.

Continuing Access, voir Accès perpétuel

Contrôle d'intégrité *Fixity check*. Méthode permettant de garantir l'intégrité d'un fichier et de vérifier qu'il n'a pas été altéré ou corrompu. Pendant le transfert, une Archive peut effectuer un contrôle d'intégrité pour s'assurer qu'un fichier transmis n'a pas été altéré en cours de route. Au sein de l'Archive, la vérification de la fixité est utilisée pour s'assurer que les fichiers numériques n'ont pas été altérés ou corrompus. Elle est le plus souvent réalisée en calculant des empreintes telles que MD5, SHA1 ou SHA256 pour un fichier et en les comparant à une valeur stockée. http://en.wikipedia.org/wiki/File_Fixity(en)

COPTR *Community Owned digital Preservation Tool Registry.* Registre des outils de préservation numérique appartenant à la communauté COPTR, hébergé par l'*Open Preservation Foundation*. http://coptr.digipres.org

Crawl, voir Moissonnage.

Criminalistique numérique *Digital Forensics*. Application de méthodes et d'outils techniques scientifiques à la préservation, la collecte, la validation, l'identification, l'analyse, l'interprétation, la documentation et la présentation d'informations numériques dérivées *a posteriori* de sources numériques.

D

Dark Archive est une archive qui ne peut être consultée par aucun utilisateur actuel mais qui peut être accessible à des dates futures sous réserve de la survenance d'événements spécifiques prédéfinis (« événement déclencheur »). L'accès aux données est soit limité à quelques personnes déterminées, soit complètement restreint à tous.

DCC *Digital Curation Centre* – Centre de conservation numérique. Une organisation basée au Irlande et active dans la conservation numérique. http://www.dcc.ac.uk

DDI *Data Documentation Initiative* (Initiative de documentation des données). Une norme internationale *de facto* pour les métadonnées permettant de décrire les données issues des sciences sociales, comportementales et économiques. http://www.icpsr.umich.edu/DDI

Designated Community, voir Communauté d'utilisateurs cible.

Digital Archiving, voir Archivage numérique.

Digital Forensics, voir Criminalistique numérique.

Dim Archive assure la préservation du train de bits pour le contenu ainsi que la planification et les actions de préservation numérique pour un accès perpétuel à long terme, et aussi un accès courant limité (peut-être limité aux utilisateurs sur place ou aux anciens abonnés après l'annulation, etc.)

DigCurV *Digital Curator Vocational Education Europe.* Un projet financé par la Commission européenne pour établir un cadre d'enseignement pour la formation professionnelle à la conservation numérique. https://www.digcurv.gla.ac.uk/

Digital Materials, voir Objets numériques.

Digital Preservation, voir Préservation numérique.

Digital Preservation Management Workshop and Tutorial, voir Atelier et tutoriel sur la gestion de la préservation numérique

Digital Publications, voir Publications numériques

Digitisation, voir Numérisation.

DIP *Dissemination Information Package* (Paquet d'informations diffusé). Un paquet d'informations envoyé par l'Archive à l'Utilisateur en réponse à sa requête à l'OAIS. Ce paquet provient d'un ou de plusieurs Paquets d'informations archivés (AIP) (terme OAIS).

DLF *Digital Library Federation* (Fédération des bibliothèques numériques). Une organisation basée aux États-Unis et active dans la préservation numérique. http://www.diglib.org

Documentation L'information fournie par un créateur et le dépôt d'archives qui fournit suffisamment d'informations pour établir la provenance, l'histoire et le contexte et pour permettre son utilisation par d'autres. Voir aussi Métadonnées.

Documents électroniques *Electronic Records*. Documents créés numériquement dans le cadre des activités quotidiennes de l'organisme et auxquels l'organisme a attribué un statut officiel. Il peut s'agir par exemple de documents de traitement de texte, de courriels, de bases de données ou de pages web intranet.

Documents numériques voir Documents électroniques

DOI *Digital Object Identifier* (Identificateur d'objet numérique). Une infrastructure technique et organisationnelle pour l'enregistrement et l'utilisation d'identificateurs pérennes largement utilisés dans les publications numériques et pour les données de recherche. Le système DOI a été créé par l'*International DOI Foundation* et a été adopté comme norme internationale ISO 26324 en 2012. http://www.doi.org

DPC *Digital Preservation Coalition*. Une organisation basée au Royaume-Uni et en Irlande, active dans le domaine de la préservation numérique et responsable du *Digital Preservation Handbook*. http://www.dpconline.org

DPTP *Digital Preservation Training Programme*, un cours de formation intensif géré par le centre informatique de l'Université de Londres. https://dptp.london.ac.uk/

DRAMBORA *Digital Repository Audit Methodology Based on Risk Assessment* (Méthodologie d'audit du dépôt numérique basée sur l'évaluation des risques). Un ensemble d'outils d'évaluation des risques développés par le DCC. http://www.dcc.ac.uk/resources/repository-audit-and-assessment/drambora

DROID Un outil de profilage de fichiers développé et distribué par TNA pour identifier les formats de fichiers. Basé sur PRONOM. http://www.nationalarchives.gov.uk/information-management/manage-information/policy-process/digital-continuity/file-profiling-tool-droid/

E

Electronic records, voir Documents électroniques.

Empreinte *Checksum.* Une signature numérique unique obtenue à partir d'un fichier. Elle est utilisée pour comparer les copies.

Émulation Moyen de surmonter l'obsolescence technologique du matériel et des logiciels en développant des techniques permettant d'imiter des systèmes obsolètes sur les générations d'ordinateurs futures.

Escrow, voir Séquestre

Événement déclencheur *Trigger Event*. Cette terminologie est utilisée lorsque des conditions spécifiques relatives à une publication électronique et à sa livraison continue aux utilisateurs sont remplies. Si la publication n'est plus disponible pour les utilisateurs auprès de l'éditeur ou de toute autre source pour diverses raisons, on dit qu'un événement déclencheur s'est produit. Ils peuvent mettre en route l'accès des utilisateurs via une Archive où la publication électronique peut être préservée numériquement.

F

FIAF Fédération Internationale des Archives du Film, une association regroupant les principales archives cinématographiques du monde. http://www.fiafnet.org

FIAT Fédération Internationale des Archives de Télévision, une association professionnelle pour ceux qui sont engagés dans la préservation et l'exploitation des archives de radiodiffusion. http://fiatifta.org

File format, voir Format de fichier

Fixity Check, voir Contrôle d'intégrité

Format de fichier *File format.* Un format de fichier est un moyen standard d'encodage des informations en vue de leur stockage dans un fichier informatique. Il indique à l'ordinateur comment afficher, imprimer, traiter et sauvegarder les informations. Il est dicté par le programme d'application qui a créé le fichier, et le système d'exploitation sous lequel il a été créé et stocké. Certains formats de fichier sont conçus pour des types de données très particuliers, d'autres peuvent servir de conteneur pour différents types de données. Un format de fichier particulier est souvent indiqué par une extension de nom de fichier contenant trois ou quatre lettres qui identifient le format. http://en.wikipedia.org/wiki/File_format (en) ou https://fr.wikipedia.org/wiki/Format_de_donn%C3%A9es (fr)

G

Garantie de l'information *Information Assurance*. Un aspect de la sécurité numérique, visant spécifiquement à garantir que la qualité de l'information est manifestement sauvegardée, qu'elle n'a pas été altérée ou qu'on n'y a pas accédé de manière inappropriée.

Gestion du cycle de vie *Life-cycle Management*. Les pratiques de gestion des documents ont établi la gestion du cycle de vie depuis de nombreuses années, tant pour les documents papier que pour les documents électroniques. Les principales implications pour la gestion du cycle de vie des ressources numériques, quelle que soit leur forme ou leur fonction, sont la nécessité de gérer activement la ressource à chaque étape de son cycle de vie, de reconnaître les interdépendances entre chaque étape et de commencer les activités de préservation dès que possible. Cela représente une différence majeure avec la plupart des méthodes de conservation traditionnelles, où la gestion est largement passive jusqu'à ce qu'un travail de conservation détaillé soit nécessaire, généralement plusieurs années après la création et n'impliquant que rarement, voire jamais, le créateur. Il existe un cycle de vie actif et interconnecté des ressources numériques, ce qui a incité de nombreuses personnes à promouvoir le terme « continuum » pour le distinguer du flux plus traditionnel et linéaire du cycle de vie des objets analogiques traditionnels. Nous avons utilisé le terme « cycle de vie » pour appliquer ce concept proactif de gestion de la préservation des objets numériques.

GIF *Graphic Interchange Format*, une image qui utilise généralement une compression avec perte. http://en.wikipedia.org/wiki/GIF (en) ou https://fr.wikipedia.org/wiki/Graphics_Interchange_Format (fr)

Gigabyte(GB), voir Gigaoctet

Gigaoctet (Go) Une unité d'information numérique souvent utilisée pour décrire les données ou la taille de stockage des données, équivaut à environ 1 000 mégaoctets (Mo).

GIS Geographical Information System, voir SIG

H

HTML *Hypertext Markup Language*. Un format utilisé pour présenter du texte et d'autres informations sur le *World Wide Web*. Depuis 1996, des versions de la spécification HTML sont maintenues par le *World Wide Web Consortium* (W3C). http://en.wikipedia.org/wiki/HTML (en) ou https://fr.wikipedia.org/wiki/Hypertext_Markup_Language (fr)

I

IASA *International Association of Sound and Audiovisual Archives*. Une association pour les archives qui préserve les documents sonores et audiovisuels enregistrés. http://www.iasa-web.org

IIPC *The International Internet Preservation Consortium*. Le Consortium international pour la préservation de l'Internet. http://www.netpreserve.org

Information Assurance, voir Garantie de l'information.

Ingest, voir Versement

InterPARES project, voir Projet InterPARES

ISO Organisation internationale de normalisation. https://www.iso.org/fr/home.html

J

Jhove2 Un outil de caractérisation des objets numériques [NdT : il est à noter que cet outil n'est plus maintenu]. La caractérisation comprend quatre éléments : l'identification du format de l'objet ;

la validation de la conformité de l'objet aux normes techniques de son format ; l'extraction des métadonnées techniques de l'objet ; et l'évaluation de l'opportunité d'accepter l'objet dans un dépôt, sur la base des politiques définies par le conservateur. https://bitbucket.org/jhove2/main/wiki/Home

JPEG *Joint Photographic Experts Group*, un comité qui supervise les normes internationales pour la compression et le traitement des photographies numériques. La majorité des formats JPEG sont avec perte. http://www.jpeg.org/

JPEG 2000 Une révision du format JPEG qui peut utiliser une compression sans perte.

K

Kilobyte (KB), voir Kilo-octet

Kilo-octet (Ko) Une unité d'information numérique souvent utilisée pour décrire des données ou la taille de stockage des données, équivaut à 1 000 octets en Système International ou traditionnellement à 1 024 octets

L

Life-cycle Management, voir Gestion du cycle de vie

Lossless Compression, voir Compression sans perte

Lossy Compression, voir Compression avec pertes

LOTAR *Long Term Archiving and Retrieval*. Norme de préservation numérique pour les modèles CAO 3D et les informations de gestion des données de produits, développée par LOTAR International, un consortium industriel d'entreprises aérospatiales et de défense des États-Unis et d'Europe. http://www.lotar-international.org

M

Megabyte (MB), voir Mégaoctet

Mégaoctet (Mo) Une unité d'information numérique souvent utilisée pour décrire les données ou la taille de stockage des données, équivaut à environ 1 000 kilo-octets (Ko).

Métadonnées Informations qui décrivent les aspects importants d'une ressource. Jusqu'à présent, la plupart des discussions ont eu tendance à mettre l'accent sur les métadonnées dans un objectif de découverte des ressources. Le présent manuel met l'accent sur les métadonnées nécessaires pour gérer et préserver avec succès les objets numériques au fil du temps et qui permettront de garantir que les informations contextuelles, historiques et techniques essentielles sont préservées en même temps que l'objet numérique. Le dictionnaire de données PREMIS pour la préservation des métadonnées est devenu une norme de facto essentielle dans le domaine de la préservation numérique.

METS *Metadata Encoding and Transmission Standard*, un standard pour la présentation des métadonnées en utilisant le XML. http://www.loc.gov/standards/mets/

Migration Un moyen de surmonter l'obsolescence technologique en transférant les ressources numériques d'une génération de matériel/logiciel à la suivante. L'objectif de la migration est de

préserver le contenu intellectuel des objets numériques et de permettre aux clients de les récupérer, de les afficher et de les utiliser d'une autre manière face à une technologie en constante évolution. La migration diffère du rafraîchissement des supports de stockage dans la mesure où il n'est pas toujours possible de faire une copie numérique exacte ou de reproduire les caractéristiques et l'apparence d'origine tout en maintenant la compatibilité de la ressource avec la nouvelle génération de technologie.

MIME *Multipurpose Internet Mail Extensions*. Un protocole permettant d'inclure des informations non ASCII dans les messages électroniques. Les logiciels comprennent généralement des interpréteurs qui convertissent le contenu MIME de et vers son format natif, si nécessaire. http://en.wikipedia.org/wiki/MIME (en) ou https://fr.wikipedia.org/wiki/Multipurpose_Internet_Mail_Extensions (fr)

Moissonnage *Crawl*. Action de naviguer sur le web automatiquement et méthodiquement pour indexer ou télécharger du contenu et d'autres données sur le web. Le logiciel qui permet d'effectuer cette opération est souvent appelé « crawler ».

MPEG *Moving Picture Experts Group*. Un comité responsable de l'élaboration de normes internationales pour la compression, la décompression, le traitement et la représentation codée des images animées, du son et de leur combinaison. https://mpeg.chiariglione.org/

N

Nativement numérique *Born-Digital*. Les objets numériques nés sous forme numérique qui ne sont pas destinés à avoir un équivalent analogique, que ce soit comme source d'origine ou comme résultat de la conversion à une forme analogique. Ce terme a été utilisé dans le manuel pour les différencier 1) des objets numériques qui ont été créés à la suite de la conversion d'originaux analogiques et 2) des objets numériques qui peuvent provenir d'une source numérique mais qui ont été imprimés sur papier, par exemple certains documents électroniques.

NCDD *The Netherlands Coalition for Digital Preservation* Coalition néerlandaise pour la conservation numérique. http://www.ncdd.nl/en/

NDSA *National Digital Stewardship Alliance*, une organisation basée aux États-Unis et active dans la préservation numérique. http://www.digitalpreservation.gov/ndsa/

NESTOR *Network of Expertise in long-term STORage and accessibility of digital ressources in Germany*. Le réseau de compétence allemand pour la préservation numérique. https://www.langzeitarchivierung.de/Webs/nestor/EN/Home/home_node.html

Numérisation *Digitalisation***.** Processus de création de fichiers numériques par numérisation ou par toute autre conversion d'objet analogique. La copie numérique qui en résulte, ou le substitut numérique, serait alors classé comme objet numérique et serait ensuite soumis aux mêmes grands défis liés à la préservation de l'accès à cet objet, en tant qu'objet « nativement numérique ».

O

OAIS *Open Archival Information System* (Système d'information archivistique ouvert). Une Archive, composée d'une organisation, qui peut faire partie d'une organisation plus large, de personnes et de systèmes, qui a accepté la responsabilité de préserver l'information et de la mettre à la disposition d'une communauté d'utilisateurs cible. Elle répond à un ensemble de responsabilités,

telles que définies dans la section 4 de la norme OAIS, qui permet de distinguer une Archive OAIS des autres utilisations du terme « archive ». Le terme « ouvert » dans OAIS est utilisé pour signifier que les normes OAIS sont développées dans des forums ouverts, et il n'implique pas que l'accès à l'Archive soit illimité. L'abréviation OAIS est également utilisée couramment pour faire référence au modèle de référence du système d'information des archives ouvertes qui a défini le terme. La norme est un cadre conceptuel décrivant l'environnement, les composants fonctionnels et les objets d'information associés à un système responsable de la préservation à long terme. En tant que modèle de référence, son objectif principal est de fournir un ensemble commun de concepts et de définitions qui peuvent aider à la discussion entre les secteurs et les groupes professionnels et faciliter la spécification des archives et des systèmes de préservation numérique. Il comporte un ensemble très élémentaire d'exigences de conformité qui doivent être considérées comme minimalistes. L'OAIS a été approuvé pour la première fois en tant que norme ISO 14721 en 2002 et une deuxième édition a été publiée en 2012. Bien qu'il ait été produit sous la direction du Comité consultatif pour les systèmes de données spatiales (CCSDS), il a bénéficié d'une contribution importante des bibliothèques et des archives.

Objets numériques *Digital Materials.* Terme général englobant les substituts numériques créés à la suite de la conversion d'objets analogiques sous forme numérique (numérisation), et « nativement numérique » pour lequel il n'y a jamais eu et n'est jamais destiné à être un équivalent analogique, et les enregistrements numériques.

Octet (o) *Byte (B)* Unité d'information numérique qui se compose de huit bits. Historiquement, l'octet était le nombre de bits utilisés pour coder un seul caractère de texte dans un ordinateur et c'est pour cette raison qu'il est la plus petite unité de mémoire adressable dans de nombreuses architectures informatiques.

OPF *Open Preservation Foundation*. Anciennement *Open Planets Foundation*. http://openpreservation.org

P

PAIMAS *Space Data and Information Transfer Systems – Producer-Archive Interface – Methodology Abstract Standard*. Cette norme ISO 20652:2006 couvre les premières étapes du processus de versement défini par le modèle de référence OAIS. http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=39577

PDF *Portable Document Format*. Un ensemble de formats et de normes ouvertes géré par l'Organisation internationale de normalisation pour la production et le partage de documents électroniques, développé à l'origine par Adobe Systems. Le format original de description de page a été élaboré au fil des versions successives pour permettre l'intégration d'objets complexes tels que des fichiers d'images, de sons et d'images animées, des hyperliens, des métadonnées XML intégrées et des formulaires actualisables. Les spécifications des différentes versions et profils du format sont désormais gérées par l'Organisation internationale de normalisation. http://www.adobe.com/uk/products/acrobat/adobepdf.html

PDF/A Versions de la norme PDF destinées à l'archivage. http://www.aiim.org/Research-and-Publications/Standards/Committees/PDFA

PDI *Preservation Description Information* – Information de pérennisation. L'information nécessaire à la préservation adéquate de l'information sur le contenu et qui peut être classée comme suit : provenance, référence, fixité, contexte et information sur les droits d'accès (terme OAIS).

Perpetual Access, voir Accès perpétuel

Petabyte (PB), voir Petaoctet.

Petaoctet (Po) Une unité d'information numérique souvent utilisée pour décrire les données ou la taille de stockage des données, équivaut à environ 1 000 téraoctets (To).

PIN Pérennisation des Informations Numériques, le groupe d'intérêt national français pour la préservation numérique. http://pin.association-aristote.fr/doku.php

Post-cancellation Access, voir Accès perpétuel

PREMIS *Preservation Metadata: Implementation Strategies* (Stratégies de mise en œuvre des Métadonnées de préservation). Une norme de facto pour les métadonnées de préservation numérique. http://www.loc.gov/standards/premis/

Préservation du train de bits *Bit Preservation*. Un terme utilisé pour désigner un niveau très élémentaire de préservation de l'objet numérique telle qu'il a été soumis (littéralement, préservation des octets formant cet objet numérique). Elle peut inclure la conservation de copies de sauvegarde sur site et hors site, le contrôle des virus, le contrôle de la fixité et le rafraîchissement périodique des nouveaux supports de stockage. La préservation du train de bits n'est pas une préservation numérique, mais elle fournit un élément de base pour l'ensemble plus complet de pratiques et de processus de préservation numérique qui garantissent la survie du contenu numérique ainsi que sa facilité d'utilisation, d'affichage, de contexte et d'interprétation dans le temps.

Préservation numérique *Digital Preservation*. Se réfère à la série d'activités gérées nécessaires pour garantir un accès continu aux objets numériques aussi longtemps que nécessaire. La préservation numérique est définie de manière très large pour les besoins de cette étude et fait référence à toutes les actions nécessaires pour maintenir l'accès aux objets numériques au-delà des limites de la défaillance des médias ou des changements technologiques et organisationnels. Ces objets peuvent être des documents créés au cours des activités quotidiennes d'une organisation, des objets « nés numériques » créés dans un but spécifique (par exemple, des ressources pédagogiques) ou les produits de projets de numérisation. Ce manuel exclut spécifiquement l'utilisation potentielle de la technologie numérique pour préserver les artefacts originaux par le biais de la numérisation. Voir également la définition de la numérisation ci-dessous.

- **Préservation à court terme** – Accès aux objets numériques soit pendant une période de temps définie pendant laquelle l'utilisation est prévue mais qui ne s'étend pas au-delà du futur prévisible et/soit jusqu'à ce qu'il devienne inaccessible en raison des changements technologiques.
- **Préservation à moyen terme** – Accès continu aux objets numériques au-delà des changements technologiques pendant une période définie mais pas indéfiniment.
- **Préservation à long terme** – Accès continu aux objets numériques, ou au moins aux informations qu'ils contiennent, pour une durée indéterminée.

Projet InterPARES *International Research on Permanent Authentic Records in Electronic Systems*. Recherche internationale sur les enregistrements authentiques permanents dans les systèmes électroniques. http://www.interpares.org

PRONOM Une base de données de formats de fichiers, de produits logiciels et d'autres composants techniques nécessaires pour soutenir l'accès à long terme aux documents électroniques et autres objets numériques ayant une valeur culturelle, historique ou commerciale. Utilisée avec DROID. http://apps.nationalarchives.gov.uk/PRONOM/Default.aspx

Propriétés significatives *Significant properties*. Caractéristiques des objets numériques et intellectuels qui doivent être préservés dans le temps afin de garantir l'accessibilité, la facilité d'utilisation et la signification des objets et leur capacité à être acceptés comme (preuve de) ce qu'ils prétendent être. https://www.archives.gov/files/era/acera/pdf/significant-properties.pdf

PST *Personal Storage Table* est une extension de fichier pour les « entrepôts personnels » locaux, écrite par le programme Microsoft Outlook. Les fichiers PST contiennent des messages électroniques et des entrées de calendrier dans un format propriétaire mais ouvert, et ils peuvent être trouvés sur les lecteurs locaux ou en réseau des utilisateurs finaux de messagerie électronique. Plusieurs outils peuvent lire et migrer les fichiers PST vers d'autres formats. http://en.wikipedia.org/wiki/Personal_Storage_Table (en) ou https://fr.wikipedia.org/wiki/.pst (fr)

Publications numériques *Digital Publications.* Des objets « nativement numériques » mis à la disposition du public et soit mis à disposition, soit distribués gratuitement ou contre rémunération. Il peut s'agir de publications en réseau, disponibles sur un réseau de communication ou de publications au format physique qui sont distribuées sur des supports tels que des disquettes ou des disques optiques. Elles peuvent également être statiques ou dynamiques.

Q

R

Rafraîchissement *Refreshing.* Copie de contenu informationnel d'un support de stockage vers le même support de stockage.

Reformatage *Reformatting* Copie du contenu informationnel d'un support de stockage vers un autre support de stockage (reformatage du support) ou conversion d'un format de fichier vers un autre format de fichier (reformatage du fichier).

Refreshing, voir Rafraîchissement

Ressources numériques, voir Objets numériques

S

Sandbox Containment, voir Bac à sable

Séquestre *Escrow.* Pratique juridique répandue consistant à déposer le code source d'un contenu ou d'un logiciel auprès d'une tierce partie. Le séquestre a lieu dans le cadre d'une relation

contractuelle, formalisée dans un accord de séquestre, entre au moins trois parties : le fournisseur, le client et le tiers fournissant le service de séquestre.

SGML *Standard Generalized Markup Language* (langage de balisage généralisé). Norme ISO sur la manière de spécifier un langage de balisage ou un ensemble de balises pour les documents. http://en.wikipedia.org/wiki/Standard_Generalized_Markup_Language (en) ou https://fr.wikipedia.org/wiki/Standard_Generalized_Markup_Language (fr)

SIG Système d'information géographique, un système qui traite ensemble la cartographie et les données.

Significant properties, voir Propriétés significatives

SIP *Submission Information Package* (Paquet d'informations à verser). Un paquet d'informations qui est livré par le producteur à l'OAIS pour être utilisé dans la construction ou la mise à jour d'un ou plusieurs paquets d'informations archivés (AIP) et/ou des informations descriptives associées (terme OAIS).

SMPTE *Society of Motion Picture and Television Engineers*. Une organisation professionnelle et un organisme de normalisation technique pour la télévision et le cinéma. https://www.smpte.org

T

Tabouret à trois pieds *Three-Legged Stool.* Une approche conceptuelle de la préservation numérique qui suggère qu'un programme de préservation entièrement mis en œuvre et viable aborde les questions organisationnelles, les préoccupations technologiques et les questions de financement, en les équilibrant comme un tabouret à trois pieds. Développé dans le cadre de l'atelier et du tutoriel sur la gestion de la préservation numérique.

TDR *Trusted Digital Repository*. Un dépôt numérique de confiance a été défini comme ayant « pour mission de fournir un accès fiable et à long terme aux ressources numériques gérées à la communauté désignée, maintenant et à l'avenir ». Le TDR doit comporter les sept attributs suivants : conformité au modèle de référence pour un système d'information archivistique ouvert (OAIS), responsabilité administrative, viabilité organisationnelle, viabilité financière, adéquation technologique et procédurale, sécurité du système et responsabilité procédurale. Ce concept est important, notamment en ce qui concerne la certification des entrepôts numériques.

Terabyte (TB), voir Téraoctet

Téraoctet (To) Une unité d'information numérique souvent utilisée pour décrire les données ou la taille de stockage des données, équivaut à environ 1 000 gigaoctets (Go).

Three-Legged Stool, voir Tabouret à trois pieds.

TIFF *Tagged Image File Format*, un format commun pour les images généralement sans perte. http://en.wikipedia.org/wiki/TIFF (en) ou https://fr.wikipedia.org/wiki/Tagged_Image_File_Format (fr)

TRAC *Trusted Repository Audit and Certification*, boîte à outils pour l'audit d'un entrepôt numérique. http://www.crl.edu/sites/default/files/d6/attachments/pages/trac_0.pdf

Trigger Event, voie Événement déclencheur

U

UKWA *UK Web Archive*. http://www.webarchive.org.uk/ukwa/

V

Versement *Ingest***.** Le processus de transformation d'un paquet d'informations à verser (SIP) en un paquet d'informations archivé (AIP), c'est-à-dire mettre les données dans une Archive numérique (terme OAIS).

W

WARC Le format WARC (*Web ARChive*) est un format de conteneur pour les sites web archivés, également connu sous le nom d'ISO 28500:2009. Il s'agit d'une révision du format de fichier ARC de l'Internet Archive utilisé pour stocker les sites web collectés ou moissonnés sur le World Wide Web. http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=44717

WAV Le format standard d'enveloppe de fichier pour l'audio ; voir BWF (*Broadcast WAV Format*) pour la variante professionnelle. http://en.wikipedia.org/wiki/WAV (fr) ou https://fr.wikipedia.org/wiki/Waveform_Audio_File_Format (fr)

Writeblockers, voir Bloqueurs d'écriture

X

XML *Extensible Markup Language*. Une norme largement utilisée (dérivée de SGML), pour représenter des informations structurées, y compris des documents, des données, des configurations, des livres et des transactions. Il est géré par le *World Wide Web Consortium* (W3C).http://www.w3.org/XML/

Y

Z