

Module 7C-Section

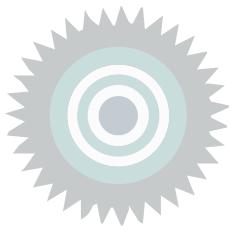
0 :Introduction à la préservation numérique

Édouard Vasseur @AIAF - PIAF

VF 02/12/2024

Table des matières

Objectifs	3
Introduction	5
1. Définition de la préservation numérique	6
1.1. Définition de « préservation »	6
1.2. Définition de « préservation numérique »	6
1.3. « Préservation numérique », « Pérennisation » ou « Archivage électronique » ?.....	7
2. La préservation numérique dans le modèle de référence OAIS	8
2.1. Description et objectifs du modèle OAIS	8
2.1.1. Qu'est-ce que le modèle OAIS ?	8
2.1.2. Le schéma descriptif du modèle OAIS	9
2.1.3. Définitions des entités fonctionnelles du modèle OAIS	10
2.2. Utilité et intérêt du modèle OAIS	10
2.3. Une approche par processus	11
3. Mythes et idées reçues sur la préservation numérique	13
3.1. Idée reçue n°1 : la préservation numérique n'est qu'une affaire de technique	13
3.2. Idée reçue n°2 : la préservation numérique est réservée aux institutions patrimoniales nationales, disposant de ressources importantes	14
3.3. Idée reçue n°3 : la préservation numérique nécessite de gros budgets	14
3.4. Idée reçue n°4 : la préservation numérique requiert des connaissances techniques avancées.....	14
3.5. Idée reçue n°5 : la préservation numérique peut attendre	15
4. Plan du module 7C	16
Glossaire	17



Objectifs

Description du module :

La préservation des documents d'archives sur support numérique – ce que les Québécois nomment documents technologiques – constitue désormais un enjeu quotidien des archivistes. L'archiviste dispose désormais d'un important panorama de normes, de standards, d'outils et de retours d'expérience pour lui permettre d'appréhender les documents d'archives sur support numérique et envisager leur préservation dans le temps.

Le but du module est de :

- aider à évaluer la situation en matière de préservation des documents d'archives sur support numérique ;
- permettre de concevoir et mettre en œuvre une politique de préservation des documents d'archives sur support numérique.

L'apprenant doit être en mesure de :

- apprêhender les spécificités en matière de préservation des documents d'archives sur support numérique ;
- dresser un état des lieux d'ensembles de documents d'archives sur support numérique ;
- définir et mettre en œuvre une politique de préservation des documents d'archives sur support numérique ;
- concevoir, mettre en œuvre et administrer un environnement permettant de gérer dans le temps les documents d'archives sur support numérique, quels que soient les moyens disponibles ;
- apprêhender les différentes catégories de formats de fichiers numériques ;
- savoir comment aller plus loin dans la réflexion.

Positionnement :

Ce module s'inscrit naturellement dans la chaîne archivistique. S'il se concentre sur les questions de planification de la préservation, de mise en œuvre de la préservation et de stockage des documents d'archives sur support numérique, il fournit également des éléments à prendre en compte lors de la mise en place de politiques et procédures de gouvernance de l'information et de gestion de l'archivage/gestion des documents d'activité/gestion des documents institutionnels/records management, de collecte de documents d'archives définitifs et d'accès à ceux-ci.

Il ne s'intéresse pas à la numérisation de documents d'archives sur support physique ou d'enregistrements sonores et audiovisuels sur support analogique, sauf dans le cas où l'opération de numérisation vise à substituer la version du document sur support numérique à celle sur support physique ou analogique.

Point sur le vocabulaire employé :

- Le terme “préservation” est entendu comme recouvrant « les fonctions de conservation préventive et matérielle » [Direction des Archives de France, Dictionnaire de terminologie archivistique, 2002] ;
- Sont distingués :
 - **les documents d'archives sur support physique**, où l'information est directement accessible à l'œil humain ou ne nécessite, pour le devenir, que l'emploi d'un appareil optique (projecteur) permettant de faciliter son agrandissement
 - **les documents d'archives sur support analogique**, où l'information, pour être intelligible, a absolument besoin de la médiation d'un appareil pour permettre à l'utilisateur de prendre connaissance de l'information (projecteur, lecteur, etc.) ;
 - **les documents d'archives sur support numérique**, qu'ils aient été directement produits avec des outils numériques ou soient le produit de la numérisation de documents d'archives sur support physique ou analogique. L'information, pour être intelligible, a absolument besoin de la médiation d'un environnement matériel et logiciel pour permettre à l'utilisateur de prendre connaissance de l'information ;
- “Document d'archives” est l'expression utilisée pour identifier toute information sur un support qui a besoin d'être prise en charge et conservée, soit pour sa valeur de preuve, soit pour sa valeur informationnelle, soit pour sa valeur patrimoniale ou de recherche. En fonction du contexte, l'expression pourra concerner des documents, des records ou des archives au sens anglo-saxon des termes ;
- “Service d'archives” est l'expression utilisée pour désigner toute structure ou organisme souhaitant mettre en place une politique de préservation de documents d'archives sur support numérique. Ce service d'archives peut être
 - interne à une organisation productrice et en charge de la gouvernance de l'information et de la gestion de l'archivage/gestion des documents d'activité/records management ou de la gestion d'archives intermédiaires ;
 - externe à une organisation productrice, soit qu'il s'agisse d'un prestataire de tiers archivage, soit d'un service d'archives définitif.

Les notions abordées dans ce module peuvent être complétées par :

- le module 9 - Section 2 : Numériser les documents qui présente les techniques de base de transfert de support vers le numérique
- le module 5 Gestion et traitement des archives courantes et intermédiaires

Il est vivement conseillé de prendre connaissance du module 7B Gestion des documents numériques au stade courant avant d'entamer la lecture du module présentement proposé. Certaines notions de base, activées ici, sont exposées plus longuement dans ce premier module.

Le glossaire du PIAF doit être consulté pour les définitions des termes spécifiques.



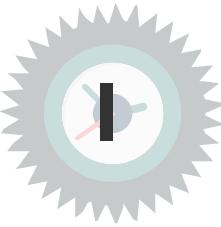
Introduction

Les documents d'archives sur support numérique tiennent une place de plus en plus importante dans notre société et dans la vie des individus et des organisations. Documents bureautiques, messages électroniques, transactions réalisées sur des applications diverses et variées disponibles en ligne ou localement (tenue d'agenda, commerce électronique, réalisation d'opérations bancaires, etc.), films et vidéos, échanges sur les réseaux sociaux, entraînements en ligne ont ainsi envahi notre quotidien.

Ce volume d'archives sur support numérique est en augmentation constante et une partie de celles-ci ont vocation à être préservées à titre temporaire ou définitif, pour leur valeur de preuve et/ou leur valeur informationnelle et leur valeur patrimoniale ou de recherche.

Comprendre les enjeux de la préservation numérique et les moyens d'y parvenir est tout l'objectif de ce module de formation.

1. Définition de la préservation numérique



1.1. Définition de « préservation »



Le terme « préservation » est d'emploi récent en français.

Inspiré de l'anglais, il désigne « les fonctions de conservation préventive et matérielle » [Direction des Archives de France, Dictionnaire de terminologie archivistique, 2002], la «fonction archivistique par laquelle est planifié l'ensemble des actions et des interventions en matière de conservation, de prévention et de restauration, et ce, en vue de protéger le patrimoine documentaire des altérations » [Charbonneau et Robert, 2001, p. 253] ou l'« ensemble d'activités hiérarchisées, planifiées et articulées de manière à assurer la durabilité de l'information. [...] Préserver] c'est [...] anticiper, mettre à l'abri ou épargner du danger de perte ou de détérioration. » [Grimard, 2009, p. 250], comme l'indique la Terminologie archivistique de base proposée aux étudiants par l'École de bibliothéconomie et des sciences de l'information de l'Université de Montréal.

1.2. Définition de « préservation numérique »

Le manuel de préservation numérique élaboré par *Digital Preservation Coalition* en 2014 et traduit en français en 2019 par un groupe d'archivistes et de bibliothécaires français, donne de l'expression « préservation numérique » la définition suivante :



« 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 [...] 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. [...] :

- 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 que l'objet 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 ;
- préservation à long terme – Accès continu aux objets numériques, ou au moins aux informations qu'ils contiennent, pour une durée indéterminée ».

1.3. « Préservation numérique », « Pérennisation » ou « Archivage électronique » ?

Même si l'expression “*préservation numérique*”^{p.17} constitue un anglicisme, nous avons préféré l'utiliser plutôt que celles de « pérennisation » et surtout d'« archivage électronique ». Cette dernière est en effet trop ambiguë en français puisqu'elle est utilisée pour traduire deux concepts différents dans la langue anglaise :

- **electronic records management**, qui définit l'ensemble des outils et procédés relatifs à la création, la réception, la gestion, l'utilisation et la mise en œuvre du sort final des documents d'archives engageants (ou documents d'activité) sur support numérique, en vue de maintenir leur valeur probante ;
- **digital preservation** en tant que telle.



Remarque

Dans le domaine du numérique, le mot « **préservation** » est concurrencé par le terme « **pérennisation** » qui désigne le « fait de maintenir l'information immédiatement compréhensible, sur le long terme, par la communauté d'utilisateurs cible, avec les preuves de son authenticité. » [Open Archive Information System OAIS (norme ISO 14721), version française, 2010] ou l'« opération préventive ou curative de régénérescence de la lisibilité d'un document imposée par l'obsolescence récurrente des technologies numériques par le biais de migrations de format et/ou du support. » [Marie-Anne Chabin, Nouveau glossaire de l'archivage, février 2010, p. 33, accessible en ligne à l'adresse suivante : https://www.marieannechabin.fr/arcateg/wp-content/uploads/2017/03/Nouveau_glossaire_de_l_archivage.pdf]. Nous avons fait le choix dans ce module de ne pas retenir ce terme, mais d'utiliser celui de « **préservation** », davantage utilisé dernièrement dans les travaux de traduction de textes sur le sujet et désormais d'emploi courant en France comme au Québec.



Fondamental

Ce choix sémantique présente un avantage certain : il a le mérite de rappeler que les documents d'archives sur support numérique sont avant tout des documents d'archives et que les questions d'évaluation, de sélection, de description, de communication, de diffusion, de valorisation et de réutilisation sont les mêmes que les documents soient sur support physique et analogique ou sur support numérique.

Si la préservation des documents d'archives sur support numérique comporte de nombreux points communs avec celle des documents d'archives sur support physique ou analogique, elle n'en présente pas moins des spécificités organisationnelles et techniques que ce module a pour vocation d'identifier et d'expliquer (voir la section 3 de ce module).

2. La préservation numérique dans le modèle de référence OAIS



Introduction

Dès que l'on parle de préservation numérique, on est rapidement confronté au modèle de référence *Open Archival Information System* – Système ouvert d'archivage d'information (*OAIS*^{p.18}).

Le modèle de référence OAIS a été élaboré à la fin des années 1990 par le Comité consultatif pour les systèmes de données spatiales (CCSDS), dont le secrétariat est assuré par l'agence spatiale américaine, la NASA.

- Une première version de ce modèle a été publiée en 1999 et a été transformée en norme internationale par l'ISO en 2002 [ISO International Standard 14721]. Ce modèle a profondément influencé toutes les réflexions en matière de préservation numérique.
- Le modèle et la norme qui en a été tirée ont fait l'objet d'une révision en 2012 [norme ISO 14721:2012 – Systèmes de transfert de données et d'informations spatiales – Système ouvert d'archivage d'information – Modèle de référence (OAIS)].

2.1. Description et objectifs du modèle OAIS

2.1.1. Qu'est-ce que le modèle OAIS ?

Le modèle de référence OAIS fournit un cadre systématique pour comprendre et mettre en œuvre les concepts et mécanismes permettant d'assurer la préservation et l'accès à long terme à des informations, notamment à ce que nous entendons par documents d'archives. Mais, plus largement, il s'adresse à tout contenu qui a vocation à être préservé sur le long terme, et à rester accessible (ouvrages de bibliothèques, autres artefacts culturels).

Conçu à l'origine pour l'information sur support numérique, il s'adapte très bien également aux informations sur support physique ou analogique.

Le modèle de référence OAIS décrit :

- les acteurs de ces opérations de préservation, réduits au nombre de trois : les fournisseurs d'informations ; les utilisateurs d'informations (ce que la norme désigne comme la « communauté désignée définie ») ; le gestionnaire d'informations ;
- les rôles et responsabilités de ces acteurs ;
- les processus et méthodes à mettre en œuvre pour préserver les informations et les rendre accessibles sur le long terme ;
- le modèle d'information à adopter pour garantir la préservation de l'information numérique et son accès à long terme, à savoir non seulement les objets contenant l'information à préserver (*Data Objects* ou objets de données), mais aussi toutes les informations utilisées pour décrire ceux-ci et pour les rendre intelligibles sur le long terme (*Representation Information* ou informations de représentation). Nous reparlerons de cet aspect dans la section 1.

2.1.2. Le schéma descriptif du modèle OAIS

Le modèle de référence OAIS est représenté sous la forme du schéma ci-dessous :

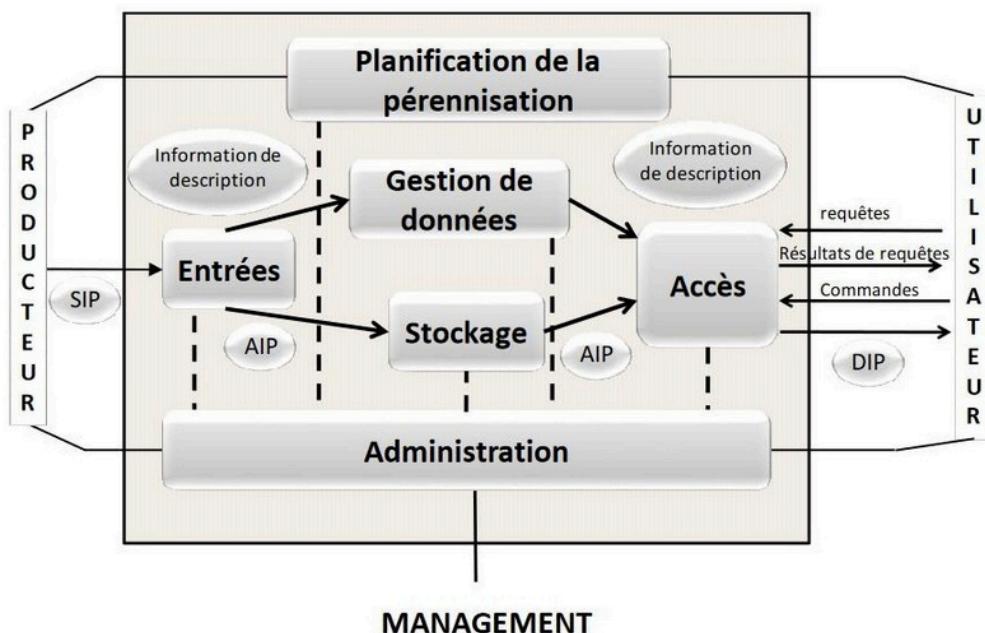


Fig.1 : Modélisation d'un OAIS (source: Modèle de référence pour un Système ouvert d'archivage d'information (OAIS), livre magenta, 2017, [https://public.ccsds.org/Pubs/650x0m2%28F%29.pdf¹](https://public.ccsds.org/Pubs/650x0m2%28F%29.pdf), p. 4-1)

- Le rectangle au centre représente ce que le modèle appelle l'« Archive », définie comme « 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 » (source: Modèle de référence pour un Système ouvert d'archivage d'information (OAIS), livre magenta, 2017, <https://public.ccsds.org/Pubs/650x0m2%28F%29.pdf>, p. 1-8). Un service d'archives constitue une « Archive », mais une bibliothèque, un musée ou une unité administrative d'une organisation privée peuvent également être considérés comme « Archive » au sens du modèle.
 - À gauche, en bas et à droite du rectangle, on retrouve les trois acteurs du modèle : les fournisseurs des informations (que le modèle désigne sous le terme de « producteur »), les utilisateurs des informations (que le modèle désigne sous le terme de « utilisateur »), le gestionnaire des informations (que le modèle désigne sous le terme de « management »). Chaque acteur est relié par des flèches aux processus et méthodes auxquels il contribue ;
 - Les processus et méthodes à mettre en œuvre pour assurer la préservation des informations sur le long terme sont au nombre de six : les « Entrées » ; la « Gestion des données » ; le « Stockage » ; l'« Accès » ; la « Planification de la pérennisation » ; l'« Administration » (voir 2.1.3. pour plus de précisions). On retrouve là les grandes fonctions mises en œuvre par un service d'archives ;
 - Le schéma offre aussi une modélisation de ce qui est échangé entre acteurs, tout au long du processus de préservation. Le modèle OAIS les désigne par l'expression de « paquets d'information » (*Information Packages*) et explique qu'ils sont de trois sortes : les paquets remis/soumis par le fournisseur au gestionnaire (les paquets d'information « soumis », *Submission Information Packages* ou *SIP* p.¹⁸) ; les paquets remis par le gestionnaire aux utilisateurs (les paquets d'information « diffusés », « disséminés », *Dissemination Information packages* ou *DIP* p.¹⁷) ; et enfin les paquets tels que gérés par le gestionnaire (les paquets d'information « archivés », *Archival Information Packages* ou *AIP*).

¹. [https://public.ccsds.org/Pubs/650x0m2\(F\).pdf](https://public.ccsds.org/Pubs/650x0m2(F).pdf)

2. [https://public.ccsds.org/Pubs/650x0m2\(F\).pdf](https://public.ccsds.org/Pubs/650x0m2(F).pdf)

2.1.3. Définitions des entités fonctionnelles du modèle OAIS

Définition des processus et méthodes à mettre en œuvre pour assurer la préservation des informations sur le long terme – ce que le modèle de référence OAIS appelle des entités fonctionnelles (source: Modèle de référence pour un Système ouvert d'archivage d'information (OAIS), livre magenta, 2017, <https://public.ccsds.org/Pubs/650x0m2%28F%29.pdf>³, p. 1-10 et 1-11)

Accès : entité fonctionnelle de l'OAIS regroupant les fonctions et services de mise à disposition des utilisateurs des fonds d'informations archivés sous la forme de Paquets d'information à diffuser ou à disséminer (DIP - *Dissemination Information Packages*).

Administration : entité fonctionnelle de l'OAIS regroupant les fonctions et services requis pour une supervision quotidienne du fonctionnement des autres entités de l'OAIS.

Entrées : entité fonctionnelle de l'OAIS regroupant les fonctions et services prenant en charge les Paquets d'informations à verser (SIP - *Submission Information Packages*) livrés par les Producteurs, préparant les Paquets d'informations archivés (AIP - *Archival Information Packages*) en vue de leur stockage, et assurant la bonne intégration dans l'OAIS de ces AIP et de leur Information de description.

Gestion des données : entité fonctionnelle de l'OAIS regroupant les fonctions et services dédiés à l'alimentation, la maintenance et la mise à disposition d'informations aussi variées que les catalogues et inventaires de ce qu'il est possible de se procurer auprès de l'entité « Stockage », les algorithmes de traitement applicables aux données récupérées, les statistiques de consultation, les éléments de facturation, les commandes par abonnement, les contrôles de sécurité, les plannings, la politique et les procédures propres à l'OAIS.

Planification de la pérennisation : entité fonctionnelle de l'OAIS qui assure les fonctions et services relatifs à la surveillance de l'environnement de l'OAIS et à la production de recommandations visant à ce que les informations stockées dans l'OAIS restent accessibles, compréhensibles et suffisamment utilisables pour la communauté d'utilisateurs cible sur le long terme, même si l'environnement informatique d'origine devient obsolète.

Stockage : entité fonctionnelle de l'OAIS regroupant les fonctions et services utilisés pour le stockage et la récupération des Paquets d'informations archivés (AIP).

2.2. Utilité et intérêt du modèle OAIS

Pourquoi ce vocabulaire et cette distinction entre les différents types d'information ? Il faut y voir plusieurs objectifs.

- Le premier est de rappeler qu'il n'y a pas que les informations à préserver qui sont manipulées lors des opérations de préservation. Il y a aussi les informations permettant de les décrire (ce que nous appellerons les métadonnées de description, désignées dans le schéma par l'expression « information de description »), les informations permettant de les gérer intellectuellement dans le temps (ce que nous appellerons les métadonnées de gestion, comme les règles de communicabilité, les droits de propriété intellectuelle applicables, les exigences imposées par les donateurs ou les dépositaires), les informations permettant de les gérer physiquement dans le temps (type de support, taille), et enfin les informations permettant d'interpréter ces informations (le dictionnaire nécessaire pour comprendre le vocabulaire utilisé, par exemple).
- Le second est de suggérer qu'il n'y a pas strictement équivalence entre ce qui est remis par le fournisseur, ce qui est géré par le service d'archives, et ce qui est remis aux utilisateurs. Pour des questions de préservation, un lot d'archives peut être réparti entre différents magasins de stockage, d'autant plus si le lot remis mélange informations sur support physique, informations sur support analogique et informations sur support numérique. Quant aux utilisateurs, ils vont demander à ce que des lots correspondant à leur demande soient constitués par le service

³. [https://public.ccsds.org/Pubs/650x0m2\(F\).pdf](https://public.ccsds.org/Pubs/650x0m2(F).pdf)

d'archives, indépendamment de l'origine des différentes informations ou du support de celles-ci. D'où la différence entre ce qui est pris en charge par le service d'archives, ce qui est géré par le service d'archives et ce que celui-ci remet aux utilisateurs. Au sens de l'OAIS, un SIP peut être conservé sous forme de 1 ou n AIP et un DIP peut être constitué d'archives conservées dans tout ou partie de 1 à n DIP.

Complément

Quatre types de métadonnées :

- les *métadonnées de description* p.17 (ex. service producteur, intitulé, date de création, etc.)
- les *métadonnées de gestion* p.17 (ex. durée d'utilité administrative, délai de communicabilité, titulaire des droits de propriété intellectuelle, etc.)
- les informations permettant de gérer les documents d'archives dans le temps (ex. type de support, taille des fichiers)
- les informations permettant d'interpréter les informations à préserver (ex. dictionnaire)

2.3. Une approche par processus

Dans le modèle de référence OAIS, la préservation numérique à proprement parler intervient dans tous les processus, mais principalement dans ceux de planification de la pérennisation, de stockage et d'administration. Ce sont ces trois processus qui sont au cœur de la préservation numérique.

Néanmoins, des actions concourant à la préservation numérique ou incluant des actions de préservation numérique peuvent intervenir dans les autres processus, à savoir les entrées, l'accès et la gestion des données. Simplement, ces processus font intervenir d'autres actions qui ne relèvent pas de la préservation et qui sont valables quel que soit le support des informations.

Exemple

Le processus des entrées fait appel aux questions d'*évaluation* p.17, de sélection, de *tri* p.18, applicables à toutes les informations. Que ces documents d'archives soient sur support physique, analogique ou numérique, l'évaluation est la même et doit d'abord prendre en compte leur *valeur probante* p.18, leur valeur informationnelle et leur *valeur patrimoniale* p.18 ou de recherche. Les spécificités techniques des documents d'archives sur support numérique doivent cependant être prises en compte dans ces actions, même si elles n'en constituent pas le cœur.

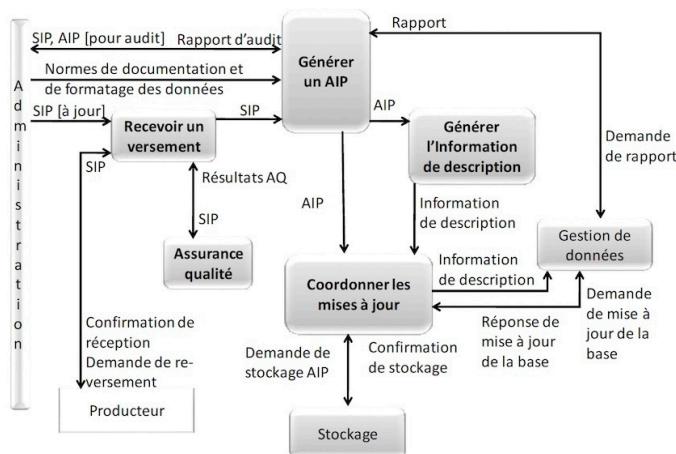
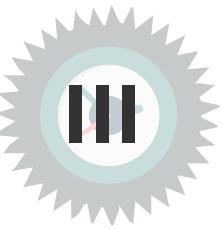


Fig.2 : Fonctions de l'Entité « Entrées » (source : Modèle de référence pour un Système ouvert d'archivage d'information (OAIS), livre magenta, pratique recommandée par le comité consultatif pour les systèmes de données spatiales, 2e édition, 2017, page 4-6, <https://public.ccsds.org/Pubs/650x0m2%28F%29.pdf>)

Conclusion

Conçu comme un ensemble de concepts et de définitions permettant de faciliter les échanges entre fournisseurs, utilisateurs et gestionnaires, quelle que soit leur culture d'origine (gestion de l'archivage/records management, gestion des archives définitives, spécialistes de l'informatique, chercheurs, grand public, etc.), le modèle de référence OAIS vise à faciliter les discussions avec les ingénieurs et autres informaticiens, d'où son caractère très générique et le vocabulaire « décontextualisé » qu'il emploie. Il n'en reste pas moins un cadre de haut niveau, parfois complexe à comprendre, alors qu'il décrit en réalité des choses simples, mises en œuvre chaque jour par les services d'archives.

3. Mythes et idées reçues sur la préservation numérique



Introduction

Avant d'entamer le module à proprement parler, il convient de revenir sur quelques idées reçues concernant la préservation numérique.

3.1. Idée reçue n°1 : la préservation numérique n'est qu'une affaire de technique

FAUX

Certes, on ne peut pas nier que les questions techniques sont essentielles dans la préservation numérique, mais elles ne sont pas les seules à devoir être prises en compte.

La préservation numérique est aussi une question :

- **archivistique** : il est important de **garder un point de vue « métier »** sur les informations numériques et de bien identifier ce qui doit être archivé et les problèmes que posent les spécificités de l'information numérique en matière de représentation et de manipulation des informations (multiplication des versions de fichiers, caractère dynamique des ressources). L'archiviste doit cependant rester lui-même face à des documents d'archives sur support numérique et mettre en œuvre les savoirs (diplomatique) et savoir-faire (évaluation, sélection, tri, description) qu'il met déjà en œuvre pour les documents sur supports physique et analogique ;
- **organisationnelle** : il faut définir des rôles et responsabilités (qui évoluent dans le temps) ; identifier ce qui peut être fait de manière internalisée par les services d'archives ou externalisée chez des prestataires de services (publics comme privés) ; identifier les compétences et les ressources disponibles ; mettre en œuvre des expérimentations avant de prendre des décisions, mais aussi formaliser les retours d'expérience et en tirer les meilleures leçons possibles à l'avenir ;
- **juridique** : au-delà des questions bien connues des archivistes qui ne changent pas pour les documents d'archives sur support numérique (communicabilité, accès des citoyens à l'information), la préservation numérique met en jeu des questions inédites de propriété intellectuelle (droit d'auteur, droit des logiciels et bases de données par exemple, mais aussi capacité à intervenir sur les œuvres de l'esprit pour assurer leur préservation, avec l'accord des ayants droit) et accentue certaines problématiques déjà présentes pour les documents d'archives sur supports physique et analogique (protection de la vie privée et des données à caractère personnel) ;
- **financière** : les compétences et les infrastructures nécessaires à la préservation des informations sur support numérique ont un coût qu'il convient d'évaluer. De même qu'il peut être trop coûteux de trier finement le contenu d'une volumineuse série de dossiers sur support physique, de même la mise en œuvre d'actions de préservation (migration de format par exemple) sur un ensemble d'archives sur support numérique pourra être estimée trop coûteuse en temps, en argent et en ressources informatiques ;

- **écologique** ; les enjeux du développement durable sont très présents avec le numérique du fait de la consommation énergétique et de l'impact environnemental des data centers.

La gestion des risques est au centre de la préservation numérique. Elle est le prisme essentiel, à la fois pour identifier les besoins et les priorités en matière de préservation numérique, mais aussi pour évaluer l'intérêt à mettre en œuvre ou non des actions de préservation numérique. L'évaluation des risques est un processus continu, itératif, à toutes les étapes de la gestion du *cycl^e de vie*^{p.17} des documents d'archives.

La gestion des risques informationnels est abordée plus en détail dans le module 7B, section 8, chapitre 3.

3.2. Idée reçue n°2 : la préservation numérique est réservée aux institutions patrimoniales nationales, disposant de ressources importantes

FAUX également

La préservation numérique peut être prise en charge par des institutions de taille moyenne voire petite. Simplement, les moyens mis en œuvre doivent être adaptés à l'environnement organisationnel, financier et technique des organisations concernées. Des outils et des services existent désormais et sont à la disposition des archivistes, quels que soient la taille et les moyens dont disposent les services d'archives.

En revanche, comme pour les documents d'archives sur supports physique et analogique, l'archiviste d'une petite structure n'aura pas nécessairement toutes les compétences nécessaires pour appréhender toutes les catégories de formats de fichiers numériques. De la même manière, les archivistes ne sont pas tous compétents en matière de traitement des plans sur calque, de restauration des parchemins ou des reliures anciennes. L'archiviste des petites et moyennes structures aura donc intérêt à s'appuyer sur les compétences des archivistes des plus grandes organisations, sur des experts appartenant à des sociétés de conseils privées.

3.3. Idée reçue n°3 : la préservation numérique nécessite de gros budgets

FAUX encore une fois

Certes, les grandes institutions nationales investissent des sommes importantes pour concevoir des outils adaptés à leurs besoins et à la volumétrie des documents d'archives sur support numérique qu'elles doivent prendre en charge. Mais il est également possible de mettre en œuvre des actions concrètes à des tarifs tout à fait raisonnables.

3.4. Idée reçue n°4 : la préservation numérique requiert des connaissances techniques avancées

VRAI et FAUX

C'est vrai dans la mesure où l'archiviste, pour mettre en œuvre certaines actions de préservation numérique, doit sortir de sa zone de confort, dialoguer avec des experts techniques et recourir à leurs connaissances et compétences spécialisées.

Mais c'est faux dans la mesure où les documents d'archives sur support numérique restent des documents comme les autres. Et surtout, les ressources disponibles aujourd'hui pour aider l'archiviste à prendre en mains les documents d'archives sur support numérique sont beaucoup plus importantes qu'il y a quelques années - manuels, revues, formations en ligne (dont celle-ci), tutoriels, webinaires, etc.

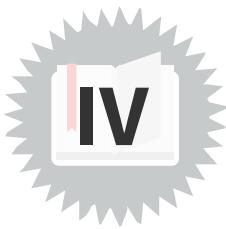
La section 6 de ce module vous aidera à identifier les ressources utiles.

3.5. Idée reçue n°5 : la préservation numérique peut attendre

FAUX.

La rapidité avec laquelle évoluent les matériels et les logiciels nécessite d'intervenir le plus tôt possible pour mettre en œuvre des opérations de préservation numérique et garantir ainsi l'accès à long terme aux documents d'archives sur support numérique.

C'est d'autant plus vrai si les documents d'archives sur support numérique ont fait l'objet d'un codage avancé et qu'aucune documentation n'est disponible. Dans ce cas, faute d'intervenir le plus tôt possible, le risque est que les documents soient perdus à jamais ...



4. Plan du module 7C

Ce module sur la préservation numérique se décompose en six sections :

- la section 1 explique les caractéristiques des documents d'archives sur support numérique et les enjeux spécifiques de leur préservation dans le temps ;
- la section 2 présente des méthodes et des outils pour établir un diagnostic de la mise en œuvre de la préservation numérique dans une institution, ainsi que pour caractériser les documents d'archives sur supports numériques d'ores et déjà pris en charge ;
- la section 3 présente les méthodes et processus à mettre en œuvre pour préserver à long terme les documents d'archives sur support numérique ;
- la section 4 présente les différents types de plateformes permettant d'assurer la préservation numérique à long terme, ainsi que les processus d'audit et de certification de celles-ci ;
- la section 5 traite de la politique de préservation et des questions spécifiques posées par les différentes catégories de formats de fichiers ;
- la section 6 présente les ressources disponibles pour approfondir les différentes notions présentées.

Glossaire



Cycle de vie

Étapes de l'existence d'un document de sa création à la mise en œuvre de son sort final.

DIP (Dissemination Information Packages)

Dans le modèle conceptuel de référence OAIS⁴, le DIP [dissemination information package] désigne le paquet d'information à diffuser ou communiquer, autrement dit l'ensemble des données et documents répondant aux requêtes d'un utilisateur.

Evaluation

Procédure préalable à l'élaboration d'un échéancier de conservation visant à déterminer les valeurs directes et indirectes et le traitement final des documents.

Métadonnées de gestion

Les métadonnées de gestion ou métadonnées administratives comprennent les métadonnées d'identification (cote par exemple), les métadonnées d'intégrité (signature électronique, horodatage, empreinte par exemple), les métadonnées décrivant les droits (gérant l'accès, la communicabilité, la réutilisation).

Métadonnées descriptives

Les métadonnées descriptives sont les métadonnées qui servent à organiser la connaissance. Ce sont les métadonnées qui identifient, classifient, hiérarchisent l'information contenue dans l'objet numérique.

Préservation numérique

La préservation numérique « réfère à la série d'activités jugées nécessaires pour garantir un accès continu aux objets numériques aussi longtemps que nécessaire » (Digital Preservation Coalition, 2021, p. 282).

⁴. <https://hub.mintika.fr/glossary/oaist-open-archival-information-system/>

SIP (Submission Information Packages)

Dans le modèle conceptuel de référence OAIS⁵, le SIP⁶ [submission information package] désigne le paquet d'information à verser, autrement dit l'ensemble des données et documents composant un versement accompagnés de leur bordereau de versement⁷.

Dans le contexte des archives publiques françaises, le SIP⁸ se compose le plus souvent d'un bordereau de versement (manifest⁹) exprimé en xml¹⁰ et conforme au SEDA¹¹ et des objets-données¹² réunis dans un dossier Content.

Source : définition proposée par mintika.<https://hub.mintika.fr/glossary/sip-submission-information-package/> Consulté le 19.11.2024

Système ouvert d'archivage d'information (Open Archival Information System - OAIS)

Modèle conceptuel de référence destiné à la gestion, à l'archivage et à la préservation à long terme de documents et données. Le modèle OAIS [Open Archival Information System] décrit les fonctions, les responsabilités et l'organisation d'un système qui voudrait préserver de l'information, en particulier des données numériques, sur le long terme, pour en garantir l'accès à une communauté d'utilisateurs identifiés.

Source : définition adaptée par mintika de la page OAIS de Wikipedia.<https://hub.mintika.fr/glossary/oais-open-archival-information-system/> Consulté le 19.11.2024

Tri

Terme générique désignant toute opération consistant à séparer, dans un ensemble de documents, ceux qui doivent être conservés de ceux qui sont destinés à être détruits (échantillonnage, sélection etc.).

Valeur patrimoniale

Valeur d'un document au regard d'une part, de sa présentation matérielle, de sa forme et/ou de son contenu et, d'autre part, de la mémoire d'un pays, qui détermine sa conservation définitive (Archives de France, Dictionnaire de terminologie archivistique, 2002).

Valeur probante

Valeur liée à une information ou un document prouvant une opération dont il peut être démontré qu'elle a été créée dans le cadre normal de la conduite de l'activité de l'organisme et qu'elle est intacte et complète. Pour revêtir cette valeur probante, les documents doivent rassembler les caractéristiques suivantes : authenticité, fiabilité et intégrité (définition adaptée d'ISO 30300:2011).

5. <https://hub.mintika.fr/glossary/oais-open-archival-information-system/>

6. <https://hub.mintika.fr/glossary/sip-submission-information-package/>

7. <https://hub.mintika.fr/glossary/bordereau-de-versement/>

8. <https://hub.mintika.fr/glossary/sip-submission-information-package/>

9. <https://hub.mintika.fr/glossary/manifest/>

10. <https://hub.mintika.fr/glossary/xml/>

11. <https://hub.mintika.fr/glossary/seda/>

12. <https://hub.mintika.fr/glossary/objet-de-donnees/>