

Exigences de base pour les archives numériques à long terme

1 Introduction

Le présent document définit les exigences que les archives numériques à long terme doivent respecter selon les connaissances et les bonnes pratiques actuelles. En ce qui concerne les bases et la terminologie, il suit le modèle de référence OAIS (ISO 14721), la structure, quant à elle, est basée sur le catalogue de critères nestor et le contenu s'inspire en outre des différentes approches internationales pour la certification d'archives numériques à long terme.

Les membres du CECO aspirent à respecter ces exigences de manière transparente dans la mise en place et l'exploitation de leurs archives numériques à long terme.

2 Définition

Les archives numériques à long terme ont pour objectif de faire en sorte que les documents électroniques remis à des Archives, qu'ils soient d'origine publique ou privée, restent durablement compréhensibles, mais aussi de garantir leur authenticité et de les rendre accessibles. Par durablement, on entend une durée illimitée, au minimum plusieurs générations de matériel et de logiciels informatiques.

Les principes d'archivage établis s'appliquent bien entendu aussi aux documents électroniques. Ces derniers sont soumis aux mêmes règles et aux mêmes exigences que les documents analogiques. Les archives numériques à long terme doivent en outre respecter une série d'exigences spécifiques d'ordre technique et organisationnel afin de pouvoir assumer leur tâche de préservation de l'information. Ces exigences se répartissent en trois groupes : objectifs et organisation, gestion des documents d'archives numériques, infrastructure et sécurité.

3 Les exigences pour les archives numériques à long terme

3.1 Objectifs et organisation

Dans la présente section, les exigences se rapportent au cadre organisationnel au sein duquel s'effectue l'archivage numériques à long terme, à savoir la fixation des objectifs, les conditions juridiques, les ressources financières et en personnel ainsi que la forme de l'organisation.

Bases : les dispositions légales et contractuelles sont respectées. Une stratégie (situation initiale, objectifs, tâches) et une variante de mise en œuvre pour les archives numériques à long terme sont définies et approuvées par l'instance compétente.

Ressources et ancrage institutionnel : les archives numériques à long terme demandent des ressources personnelles et financières à long terme. Elles doivent donc être financées par le budget ordinaire.

Cadre (policy) : il convient de définir des processus et des modèles de données pour chaque étape de l'archivage numériques à long terme (transmission, conservation, description, utilisation) et de déterminer des mécanismes de contrôle de ces tâches. Processus et modèles de données sont en constante mutation et doivent être régulièrement adaptés.

3.2 Gestion des documents d'archives numériques à long terme

Les exigences de cette section se rapportent à la gestion des documents d'archives numériques durant le cycle de vie de l'objet stocké dans les archives numériques à long terme. Elles se basent sur les tâches centrales définies par le modèle de référence OAIS : prise en charge (ingest), catalogage (data management), stockage (archival storage) y compris conservation de l'information donnée par la séquence de bits (bitstream preservation), utilisation (access) ainsi que les processus de soutien qui sont l'administration et la planification des mesures de conservation à long terme (preservation planning).

Chaque tâche particulière s'étend aussi bien aux données primaires numériques qu'à toutes les métadonnées nécessaires à la compréhension.

Paquet d'informations : lors de l'enregistrement des données primaires, on enregistre également un « paquet » de métadonnées descriptives et techniques (SIP, AIP). Il s'agit d'assurer de manière appropriée le lien entre données primaires et métadonnées, mais également de vérifier que les métadonnées correspondent à la structure définie. Les données représentant les documents électroniques archivés doivent en outre être enregistrées dans des formats archivables. Enfin, les archives sont clairement identifiables.

Transmission : la transmission de documents aux Archives s'effectue selon un processus documenté (*entrées*). Ce dernier permet de déterminer de manière transparente sous quelle forme et avec quelles métadonnées les documents sont pris en charge (SIP).

Bitstream preservation : les documents électroniques archivés sont classés dans une infrastructure mémoire répondant au moins aux exigences d'archivage suivantes : trois copies à deux endroits différents, contrôles réguliers de l'intégrité des données, accès réglementé aux données (*stockage*).

Aides à la recherche : Les documents numériques archivés sont répertoriés de sorte à pouvoir être retrouvés avec les moyens appropriés que ce soit par un « paquet » minimal d'informations descriptives ou par une recherche en plein texte et l'accès aux métadonnées inhérentes du document numérique. Le catalogage doit s'effectuer déjà lors du processus de prise en charge depuis le SIP.

Utilisation : les archives numériques à long terme offrent aux utilisateurs l'accès (consultation) aux documents électroniques archivés, conformément au cadre juridique (*accès*). Elles déterminent les processus et les formats d'accès correspondants (*DIP*).

Planification de la pérennisation : les archives numériques à long terme planifient et mettent en œuvre les mesures appropriées visant à garantir la lisibilité durable des documents électroniques archivés.

3.3 Infrastructure et sécurité

Les exigences de cette section se rapportent aux aspects techniques de l'ensemble du système et à la sécurité des informations.

Infrastructure et sécurité : l'infrastructure garantit la protection des archives numériques à long terme et des objets numériques qu'elles abritent. Outre les exigences découlant de la gestion des objets, l'infrastructure informatique remplit les exigences de sécurité prévues par le concept de sécurité informatique

Concept de sécurité informatique : le concept de sécurité a été défini et comprend notamment les autorisations d'accès et le compte-rendu des accès aux objets numériques. Le personnel des archives est formé en matière de sécurité informatique.

4 Annexes

4.1 Bibliographie

American Library Association,
Association for Library Collections and Technical Services
Definitions of Digital Preservation, 2007
<http://www.ala.org/alcts/resources/preserv/defdigpres0408>

Comité Consultatif pour les Systèmes de Données Spatiales
Modèle de référence pour un Système ouvert d'archivage d'information (OAIS)
CCSDS 650.0-M-2(F), Livre magenta, 2012 (traduction française 2017)
[https://public.ccsds.org/pubs/650x0m2\(F\).pdf](https://public.ccsds.org/pubs/650x0m2(F).pdf)
(correspond à ISO 14721:2012, <https://www.iso.org/standard/57284.html>)

Consultative Committee on Space Data Systems
Audit and certification of trustworthy digital repositories
CCSDS 652.0-R-1, Magenta Book, 2011
<https://public.ccsds.org/pubs/652x0m1.pdf>
(correspond à ISO 16363:2012, <https://www.iso.org/standard/56510.html>)

DCC, DPE, nestor, CRL
Core Requirements for Digital Archives, 2007
<http://www.crl.edu/content/DigArc/metrics2/corerequirements.htm>

Direction des Archives de France
L'archivage numérique à long terme : les débuts de la maturité ?
Manuels et guides pratiques
Paris, La Documentation française, 2009

Edmunds, Rorie et al
Core trustworthy data repositories requirements, 2016
<https://doi.org/10.5281/zenodo.168411>

Nestor-Kriterien
Kriterienkatalog vertrauenswürdige digitale Langzeitarchive, Version 2
(Nestor-Materialien 8), 2008
<urn:nbn:de:0008-2008021802>

OCLC/CRL
Trustworthy Repositories Audit & Certification:
Criteria and Checklist [TRAC], 2007
http://www.crl.edu/sites/default/files/d6/attachments/pages/trac_0.pdf

4.2 Versions

Version	Date	Commentaire
1.0	09.2009	Publication initiale
1.1	11.2010	Adaptation terminologie française
2.0	11.2018	Accent mis sur les archives numériques à long terme au lieu de l'archivage numérique ; Condensation de l'introduction et de la définition ; Structuration des exigences en fonction des critères nestor ; Intégration des exigences « bases », « planification de la pérennisation », « infrastructure et sécurité », « concept de sécurité informatique » ; Mise à jour de la bibliographie