



PROJET D'ARCHIVAGE PERENNE AU CINES

Guide de mise en œuvre

03 MARS 2025

CINES

950, rue de Saint Priest, 34097 Montpellier Cedex 5. France



Table des matières

1.	Introduction	4
1.1	Objet du document	4
1.2	Choix de la méthode.....	4
2.	Phase préliminaire	4
2.1	Lettre d'intention	4
2.2	Etude de faisabilité du projet	4
2.3	Lancement du projet d'archivage	5
3.	Phase de définition	5
3.1	Modèle de données.....	6
3.2	Profil d'archivage.....	7
3.3	Paquet d'archives type	7
3.4	Plan de classement.....	7
3.5	Référentiels	8
3.6	Convention d'archivage.....	8
4.	Phase de tests et de développement	9
4.1	Fiche de liaison	9
4.2	Administration fonctionnelle du SAE	10
4.3	Réalisation de tests	10
5.	Phase de validation	10
6.	Phase de production, de conservation et de traitement.....	11
7.	Phase de clôture du projet (le cas échéant)	11
8.	Conclusion.....	11

1. Introduction

1.1 Objet du document

Le but de ce document est de présenter le processus de mise en œuvre d'un projet d'archivage électronique au CINES. Il s'adresse aux utilisateurs (ou services versants) de la Plateforme d'archivage du CINES (PAC) et au personnel du Centre, plus particulièrement du Département Archivage et Service aux Données (DASD).

1.2 Choix de la méthode

Pour faciliter les échanges avec les services versants, les étapes d'un projet d'archivage électronique dans le système d'archivage électronique (SAE) PAC s'inspirent de la norme PAIMAS (*Producer Archive Interface Methodology Abstract Standard*)¹. L'application de cette norme permet d'éviter des « problèmes potentiels » tels que l'absence de correspondance entre les archives reçues par le service d'archives et celles attendues, la transmission d'archives non clairement définies par le service versant, le non-respect du calendrier des versements, et la détection tardive, voire nulle, des erreurs de transfert par le service d'archives².

Cette approche permet ainsi de s'accorder sur la nature des données à archiver, la manière de les archiver, la période durant laquelle elles pourront être versées dans le SAE et d'anticiper sur les problèmes pouvant se présenter avant ou après le passage en production.

Dès lors, un projet d'archivage au CINES se divise en cinq ou six grandes phases, chacune impliquant différents acteurs et ayant une durée d'exécution variable en fonction de la complexité du projet. À chaque étape, des livrables et/ou des actions sont attendus, et l'ensemble des documents, ainsi que les comptes rendus de réunion, constituent la documentation du projet.

2. Phase préliminaire

La phase préliminaire, ou avant-projet, constitue l'étape initiale d'un projet d'archivage, précédant sa phase opérationnelle.

Elle se déroule en trois temps : d'abord, l'envoi par le demandeur d'une lettre d'intention, suivi d'une étude pour vérifier si le projet est réalisable, puis, si l'étude est concluante, la confirmation de la faisabilité et le lancement du projet.

2.1 Lettre d'intention

La lettre d'intention, adressée à la direction du CINES, est la démarche initiale pour les organismes souhaitant archiver leurs documents dans le SAE PAC. Elle indique leur intention de conclure une convention d'archivage avec le CINES, mais ne constitue pas un engagement contractuel. Seule la convention d'archivage formalise juridiquement le cadre contractuel de l'archivage.

2.2 Etude de faisabilité du projet

Après réception de la lettre d'intention, le CINES et l'organisme demandeur engagent des discussions pour évaluer la faisabilité du projet d'archivage. Cette évaluation approfondie porte sur divers

¹ ISO 20652:2006. Systèmes de transfert des informations et données spatiales — Interface entre producteur et archives — Norme pour la méthodologie concernant les abstraits

² ISO 20652:2006, p. 11.

aspects clés : la nature des informations à archiver³, les conditions de sécurité, les aspects juridiques et contractuels, le calendrier, les coûts, les risques et points critiques⁴, le statut juridique du demandeur, les formats et la durée de conservation des documents, ainsi que leur sort final.

L'évaluation intègre également les moyens financiers, humains et techniques disponibles. Un avis favorable permet de lancer le projet et d'initier les démarches contractuelles.

2.3 Lancement du projet d'archivage

Le lancement effectif du projet d'archivage débute par une réunion inaugurale entre le CINES et le l'organisme demandeur. Cette première rencontre vise principalement à initier concrètement le projet, à rappeler le contexte et à actualiser les informations depuis la lettre d'intention. Elle permet également de clarifier l'organisation, notamment les rôles et responsabilités de chaque partie, et d'établir un calendrier. Cette étape est essentielle pour garantir une coordination efficace entre les acteurs du projet.

3. Phase de définition

La phase de définition a trois principales finalités :

- Définir, entre autres, les données et métadonnées à archiver, la structure des paquets d'archives à verser, les référentiels du projet ;
- Identifier les protocoles de communication entre le système d'information (SI) du service versant et le SAE PAC, et les modalités de versement et d'accès aux données ;
- Rédiger ou compléter et signer la convention d'archivage.

Les livrables attendus sont notamment le modèle de données, le plan de classement, le paquet d'archives type (ou SIP type), le profil d'archivage et les différents référentiels (contrat d'entrée, règles de gestion, etc.)

Pour atteindre ces objectifs, le service versant doit au préalable se rapprocher des services métiers pour recueillir les informations essentielles à la modélisation des données. Il doit répondre à plusieurs questions sur la nature, l'organisation des objets à archiver, les métadonnées, etc.

Nous présentons ci-dessous une liste non exhaustive des questions à se poser.

Nature des objets à archiver
Quelles sont les données à archiver (typologies documentaires) ?
Formats
Quels sont les formats des données à archiver ?
Est-il prévu une conversion de formats de fichiers ?
Règlementation

³ S'agit-il de données personnelles, sensibles, de santé, d'archives publiques, privées ?

⁴ « Évaluer les points critiques : Le producteur et l'archiviste doivent évaluer, parmi tous les points déjà soulevés, ceux qui peuvent poser de sérieux problèmes et qui sont susceptibles d'avoir une incidence sur le projet Producteur-Archives ». PAIMAS, mai 2004, p.37.

Les archives sont-elles publiques ou privées (au sens du code du Patrimoine ⁵) ? Contiennent-elles des données à caractère personnel ou des données de santé ? Existe-il un besoin légal ?
Capacité de stockage
Quelle est la volumétrie prévisionnelle des archives ? Combien d'objets sont estimés ?
Organisation des objets à archiver
Comment sont organisés les objets à archiver ?
Qui sont les services producteurs des documents à archiver ?
Existe-t-il un plan de classement des documents ?
Comment ces objets sont-ils représentés ? S'agit-il de fichiers ou de données stockées dans une base de données ?
Une reprise de l'existant a-t-elle été prévue ?
Métadonnées
Quelles sont les métadonnées associées aux données ?
Proviennent-elles d'un système de gestion électronique de documents (GED), d'un logiciel métier ?
Gestion des archives
Quelle est la durée de conservation des données ?
Quel est le sort final envisagé pour ces archives ?
Quels sont les droits d'accès et de consultation des données ?
Cas particuliers
Existe-t-il des exceptions aux règles générales identifiées ci-dessus ?
Les données comportent-elles une signature électronique ?
Est-il prévu de mettre à jour les données ? Si oui, à quelle fréquence ?

En répondant à ces questions, le service versant pourra fournir des informations nécessaires à la phase de définition du projet d'archivage, ainsi qu'à la phase de tests.

3.1 Modèle de données

Avant de créer un profil d'archivage et un paquet d'archives à verser, il est essentiel de définir un modèle de données (MDD). Ce MDD, basé sur le Standard d'échanges des données pour l'archivage (SEDA), décrit précisément la structure et l'organisation des données à archiver. Il précise notamment les types de données (fichiers, dossiers), leurs relations (arborescence) et les règles de validation (données obligatoires). En résumé, le MDD est une maquette de l'organisation des données au sein du futur paquet d'archives.

⁵ Articles L211-4 et L211-5 du code du patrimoine.

Le modèle de données, le paquet d'archives type et le profil d'archivage sont spécifiques à une catégorie de données ou typologie documentaire.

3.2 Profil d'archivage

« Chaque document archivé sera associé à un "profil d'archivage" qui est défini comme un ensemble de règles applicables aux documents partageant les mêmes critères de confidentialité, de durée de conservation, de destruction et de droits d'accès. »⁶

Le profil d'archivage assure le contrôle qualité des données et métadonnées d'un paquet d'archives, vérifiant la conformité du bordereau de versement (manifest.xml) lors de l'intégration dans le SAE. Il valide la structure, les métadonnées et les objets d'archives selon le standard SEDA 2.1. Sa création repose sur le modèle de données, et toute modification de ce dernier requiert une mise à jour du profil. La création du profil peut se faire via l'outil Patis⁷ (format Relax NG⁸) ou un éditeur XML (format XSD⁹ également supporté)¹⁰. L'utilisation d'un profil d'archivage est obligatoire au CINES, particulièrement pour les versements automatisés.

3.3 Paquet d'archives type

Le paquet d'archives type (ou SIP type), conforme au SEDA 2.1, sert de modèle pour les archives en production. Il doit être représentatif et complet. Sa création s'effectue avec des outils comme Resip¹¹ et Octave¹², qui permettent de manipuler et constituer des SIP SEDA¹³. Ce paquet de test reproduit fidèlement le modèle de données, incluant objets, unités archivistiques hiérarchisées et métadonnées. Les premiers tests de versement manuels utilisent ce paquet pour vérifier la conformité au profil d'archivage.

3.4 Plan de classement

La mise en place d'un plan de classement est une étape primordiale, intervenant dès la phase de définition du projet. Ce plan doit répondre aux besoins à la fois métiers et archivistiques, qui peuvent diverger. Il est essentiel de définir le contenu du paquet d'archives (SIP) avant d'élaborer le plan de classement. Par exemple, dans un projet d'archivage des ordres de mission, chaque SIP pourrait représenter un ordre de mission individuel, classé dans un dossier d'ordres de mission.

Le plan de classement n'est pas figé et peut être ajusté en fonction des évolutions du fonds d'archives ou de changements significatifs. Par exemple, l'intégration d'une nouvelle collection ou le versement dans le SAE d'une nouvelle typologie documentaire peuvent nécessiter une mise à jour du plan.

⁶ https://www.piaf-archives.org/sites/default/files/bulk_media/m07s05/co/module_section05_7.html

⁷ Profil d'Archivage Simple pour le Traitement de l'Information en SEDA

⁸ Relax NG (Regular Language for XML Next Generation) est un langage de description de document XML.

⁹ <https://www.w3.org/TR/xmlschema11-1/>, consulté le 08/04/2024.

¹⁰ https://www.programmevitam.fr/vitam-doc/fr/master_7.1.x/sections/Vademecum_pa.html, consulté le 27/02/2025.

¹¹ <https://www.programmevitam.fr/pages/ressources/resip>, consulté le 08/04/2024.

¹² <https://francearchives.gouv.fr/fr/article/88482499>, consulté le 08/04/2024.

¹³ <https://www.programmevitam.fr/pages/ressources/resip>, consulté le 08/04/2024.

Notons que dans un environnement de production, la mise en place du plan de classement doit être effectuée avant le premier versement d'archives.

3.5 Référentiels

Les référentiels sont des données de référence, structurées par catégories, pouvant être génériques ou adaptées à un projet d'archivage. Ils jouent un rôle clé dans la gestion des droits d'accès, qu'ils soient attribués à des utilisateurs humains ou à des machines, et contribuent à la sécurité et à la confidentialité des archives. Un paquet d'archives déclare les référentiels dont il dépend pour être pris en charge par le SAE.

Les principaux référentiels à définir en début de projet sont :

- Le contexte applicatif¹⁴ ;
- Les services agents¹⁵ ;
- Les règles de gestion¹⁶ ;
- Le(s) contrat(s) d'entrée¹⁷ ;
- Le(s) contrat(s) d'accès¹⁸ ;
- Le(s) profil(s) d'archivage¹⁹.

Les référentiels sont essentiels pour contrôler les versements, les accès, la provenance et le cycle de vie des archives. Pour approfondir, consultez la section « Structuration des référentiels »²⁰ du vadémécum du Programme Vitam.

3.6 Convention d'archivage

La convention d'archivage officialise le partenariat entre le service versant et le CINES. Elle a pour objectif la mise en place d'un service de préservation des archives numériques déposées par le service versant. Sa signature par l'ensemble des parties permet le passage en production.

En résumé, la phase de définition doit permettre la livraison des documents suivants :

1. Modèle de données ;
2. Profil d'archivage ;
3. Paquet d'archives type ;
4. Plan de classement ;
5. Référentiels ;
6. Convention d'archivage.

Le passage à la phase suivante du projet est conditionné par la livraison d'un paquet d'archives type conforme au profil d'archivage.

¹⁴ https://www.programmevitam.fr/vitam-doc/fr/master_7.1.x/sections/gestion_habilitations.html#contexte-applicatif

¹⁵ https://www.programmevitam.fr/vitam-doc/fr/master_7.1.x/sections/Vademecum_referentiel_sa.html

¹⁶ https://www.programmevitam.fr/vitam-doc/fr/master_7.1.x/sections/Vademecum_referentiel_rg.html

¹⁷ https://www.programmevitam.fr/vitam-doc/fr/master_7.1.x/sections/gestion_habilitations.html#contrat-dentree

¹⁸ https://www.programmevitam.fr/vitam-doc/fr/master_7.1.x/sections/gestion_habilitations.html#contrat-dacces

¹⁹ https://www.programmevitam.fr/vitam-doc/fr/master_7.1.x/sections/Vademecum_pa.html

²⁰ https://www.programmevitam.fr/vitam-doc/fr/master_7.1.x/sections/Vademecum_sommaire.html, consulté le 27/02/2025.

4. Phase de tests et de développement

La phase de tests et de développement, impliquant référents projet et équipes techniques, varie en durée selon les spécificités techniques du projet. Elle comprend, le cas échéant, le développement d'applications de versement ou de consultation, ainsi que la mise en place de connecteurs entre GED, applications métier et SAE.

Les objectifs principaux de cette phase sont de :

- Valider la connexion entre le système d'information du service versant et le SAE PAC ;
- Configurer l'environnement de test du service versant dans le SAE ;
- Évaluer les fonctionnalités de versement, de recherche et d'accès.

4.1 Fiche de liaison

L'accès au SAE PAC, basé sur la solution Vitam, nécessite la communication, par le service versant, des adresses IP des postes utilisateurs. De plus, un certificat X.509²¹ signé par une Autorité de Certification (AC) et valide pendant 13 mois (conformément aux directives du CA/B Forum)²², doit être fourni. Ce certificat doit être transmis de manière sécurisée au correspondant technique du CINES. Les certificats auto-signés ne sont pas acceptés.

Le CINES fournit au service versant les informations suivantes :

- L'identifiant du contrat d'entrée permettant le versement d'archives ;
- L'identifiant du contrat d'accès permettant l'accès aux archives ;
- Le numéro de tenant (ou coffre) de recette ;
- L'identifiant du service versant ;
- L'identifiant du profil d'archivage ;
- Les ports à ouvrir en sortie du réseau.

Ces informations sont centralisées dans une fiche de liaison, qui regroupe les éléments fonctionnels et techniques nécessaires à la configuration des environnements de test et de production. Cette fiche identifie également les interlocuteurs clés (techniques et fonctionnels) du CINES et du service versant.

Le tableau ci-dessous récapitule les différentes actions à entreprendre pour établir la connexion au SAE PAC et les responsabilités de chaque partie.

Actions	Responsabilité
Génération des certificats applicatifs et envoi au CINES pour l'accès au contexte applicatif de la solution Vitam	Service versant
Fourniture des adresses IP des serveurs devant se connecter au SAE	Service versant

²¹ « X.509 est une norme spécifiant les formats pour les certificats à clé publique, les listes de révocation de certificat, les attributs de certificat, et un algorithme de validation du chemin de certification, définie par l'Union internationale des télécommunications (UIT) ». En ligne : <https://fr.wikipedia.org/wiki/X.509>

²² Cette durée a été déterminée par le Certification Authority Browser Forum ou CA/B Forum. Le CA/B Forum détermine les périodes de validité des certificats TLS/SSL à l'issue d'un processus de vote interne prévu dans son règlement. Le standard actuel de validité maximale des certificats est de 397 jours, soit 13 mois environ, et a été voté le 1er septembre 2020.

Déclaration des certificats applicatifs du service versant dans la solution Vitam	CINES
Autorisation des adresses IP des serveurs du service versant	CINES
Ouverture en entrée du port 8443 pour l'Ingest ²³ - et si pertinent, du port 8444 pour l'Access	CINES
Ouverture en sortie du port 8443 pour l'Ingest - et si pertinent, du port 8444 pour l'Access	Service versant
Paramétrage et mise à disposition du tenant de test (Bac à sable) pour la phase de tests et de développement	CINES
Test d'interrogation de l'API Ingest	Service versant
Test d'accès réseau : par la commande telnet vitam.cines.fr 8443 et 8444	Service versant

4.2 Administration fonctionnelle du SAE

Le paramétrage du SAE PAC, également appelé administration fonctionnelle, consiste à :

- Préparer la prise en charge des paquets d'archives dans le SAE en déclarant les référentiels et le plan de classement du service versant ;
- Configurer les droits d'accès des utilisateurs aux différentes fonctionnalités du système.

Le paramétrage s'applique aux environnements de test et de production. L'environnement de test sert à vérifier les fonctionnalités du système et à détecter les erreurs avant la mise en production. L'environnement de production, quant à lui, est l'endroit où les archives sont réellement conservées. Son accès est restreint et les actions sont contrôlées pour garantir l'intégrité des archives.

Pour éviter les erreurs lors du versement d'archives, l'environnement de test doit être identique à l'environnement de production. Cela permet de vérifier que les opérations fonctionnent comme prévu avant la mise en production.

4.3 Réalisation de tests

Une fois l'environnement de test configuré, le service versant peut tester les principales fonctionnalités du SAE (versements, consultation) à partir d'une interface dédiée. Il est recommandé d'effectuer ces tests uniquement avec des documents fictifs ou non confidentiels.

5. Phase de validation

La phase de validation voit la participation des référents projet, des archivistes et des équipes techniques.

La validation comprend trois étapes essentielles : s'assurer de la signature de la convention d'archivage par toutes les parties, préparer et valider l'environnement de production du service versant, et effectuer le versement inaugural d'un paquet d'archives.

La réussite de ce premier versement dans l'environnement de production permet de passer à la phase de production, de conservation et de traitement.

²³ Entrée ou versement d'archives

6. Phase de production, de conservation et de traitement

Aucun livrable n'est prévu dans cette phase, qui se caractérise principalement par :

- le versement continu et la conservation des archives dans le SAE ;
- des opérations de communication, de restitution/transfert ou d'élimination des archives ;
- la tenue de réunions régulières pour le suivi du projet ;
- la tenue d'une réunion annuelle permettant de faire le point sur l'ensemble des aspects du projet : la convention, la volumétrie, les difficultés rencontrées, la qualité de service ou les besoins d'évolution.

7. Phase de clôture du projet (le cas échéant)

Enfin, le passage à la phase de clôture du projet intervient uniquement dans le cas d'une restitution complète des archives au service qui les a produites. Autrement dit, cette phase ne s'applique pas si la restitution ne porte que sur une partie des archives conservées.

8. Conclusion

En résumé, un projet d'archivage au CINES passe par plusieurs phases. La première phase permet de réaliser une étude de faisabilité du projet. La deuxième prépare les données à archiver et la convention d'archivage. La troisième permet de tester les fonctionnalités de versement et de consultation des archives. La quatrième, conditionnée par la signature de la convention, a pour objectif de valider les tests de l'étape précédente en faisant un premier versement en production. En cas de réussite du premier versement, le projet peut passer à la cinquième phase permettant au service versant de déposer officiellement et régulièrement ses archives dans le SAE PAC pour conservation. Enfin, la phase de clôture d'un projet n'intervient que dans le cas où une des deux parties souhaite mettre fin à la convention d'archivage.