

Archivage des données

Aujourd'hui, l'essentiel de la production scientifique est créé, traité, stocké et diffusé sous forme numérique. Cependant, avec la constante évolution des technologies, la pérennisation des informations produites est devenue un enjeu majeur.

C'est pourquoi il est important de mettre en œuvre des bonnes pratiques en matière d'archivage pérenne pour assurer la conservation des données et répondre à ces 4 objectifs : rendre les documents **accessibles**, en préserver l'**intégrité**, l'**intelligibilité** et la **lisibilité**.

Connaître les risques



L'obsolescence technologique

Depuis plus de 50 ans, l'**évolution rapide des matériels** informatiques (nouvelle génération tous les 5 ans en moyenne) engendre une **obsolescence des logiciels** qui doivent sans cesse rester compatibles avec les nouveaux matériels.



La disparition de formats

On observe en permanence la disparition de formats devenus obsolètes et l'apparition de nouveaux formats d'enregistrement des données.



Le manque de documentation

Afin de préserver l'intelligibilité des données, il est indispensable de les accompagner de toutes les métadonnées nécessaires à la bonne compréhension du contexte de leur production.

Le saviez-vous ?

Il ne faut pas confondre **stockage sécurisé** (pour assurer la continuité de l'exploitation sur du court terme) et **archivage pérenne** (conservation à moyen et long terme).

Mettre en œuvre l'archivage



Quel modèle d'archivage suivre ?

Il est recommandé de suivre le modèle **OAIS** (Open Archival Information System, ISO 14721) qui fournit les bases nécessaires à l'organisation d'un système d'archivage. Ce modèle définit les acteurs, les différentes entités fonctionnelles et la façon dont elles s'articulent entre elles.



Quel cadre juridique respecter ?

Il faut s'assurer que les exigences juridiques liées aux données permettent leur archivage : **archives publiques, données sensibles, propriété intellectuelle**... Tous ces éléments peuvent avoir des conséquences sur la durée de conservation ou sur la diffusion.

Comment agir ?

Il faut conserver des copies multiples des documents tout en privilégiant les technologies les plus utilisées, libres et/ou ouvertes, et en diversifiant les supports. Il faut aussi planifier les opérations de migration de supports.

Il faut opter pour des formats dits « pérennes », dont les spécifications sont librement accessibles. Il faut suivre les recommandations sur les formats et dispositifs d'enregistrement proposées par la BnF, le CINES et la TGIR Huma-Num.

On s'assure de la bonne compréhension du document en renseignant des métadonnées descriptives, techniques et administratives. Il faut également avoir un identifiant unique pour chaque ressource.

À noter

*La norme AFNOR **NF Z42-013** présente les exigences opérationnelles et techniques à mettre en œuvre au sein du système d'archivage électronique.*

Sollicitez les personnes expertes dans ce domaine : juristes, archivistes, le DPO (Délégué à la Protection des Données), PRADA (Personnes Responsables de l'Accès aux Documents Administratifs). Consultez le Code du Patrimoine, le Code de la Propriété Intellectuelle et le portail national des Archives ([FranceArchives](https://www.francearchives.fr)).

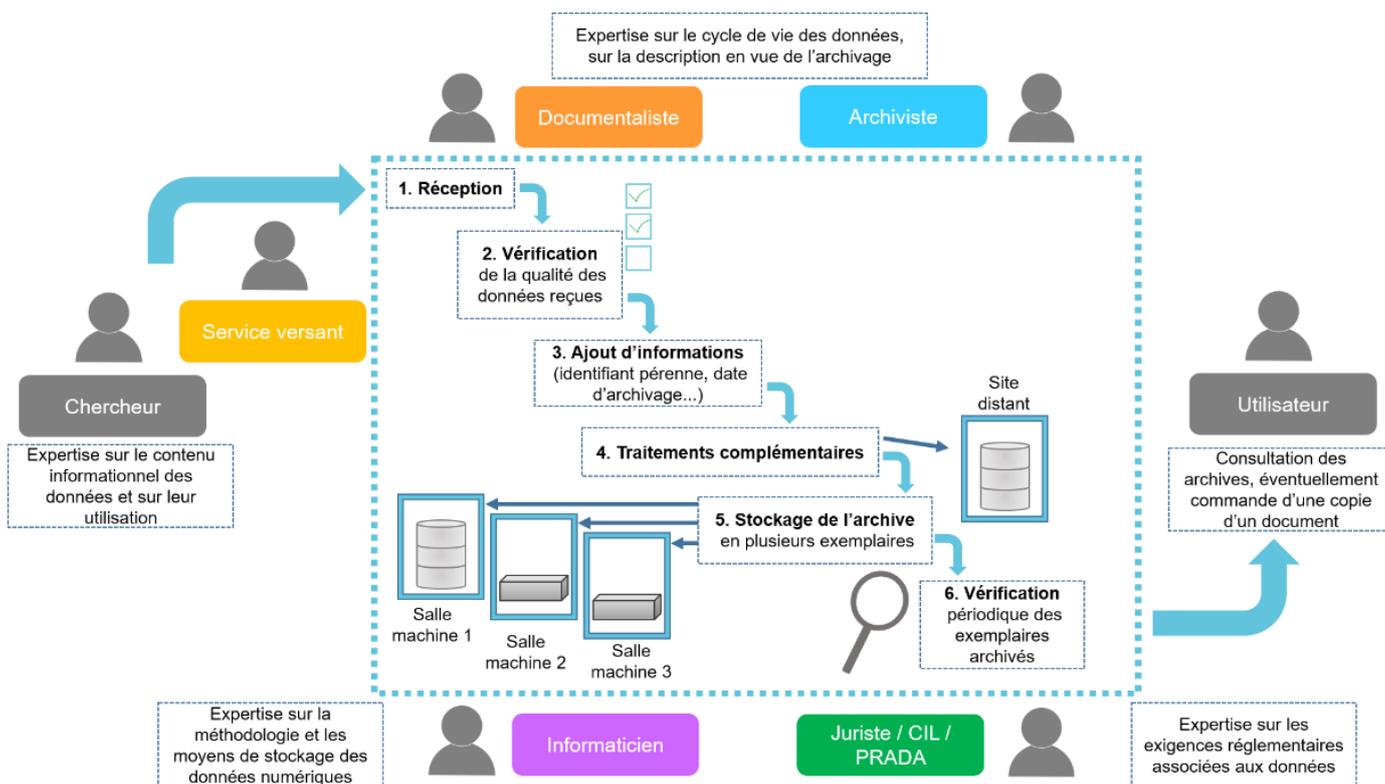
Un exemple de système d'archivage : le CINES

Le **CINES** (Centre Informatique National de l'Enseignement Supérieur) a été mandaté par le Ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation pour mettre en place une solution d'archivage électronique du patrimoine scientifique pour toute la communauté ESR (Enseignement Supérieur et Recherche). Pour remplir cette mission, le CINES propose plusieurs solutions d'archivage, dont la plateforme PAC (Plateforme d'Archivage du CINES). Ces solutions présentent plusieurs caractéristiques : elles sont mutualisées, personnalisées, sécurisées et certifiées (agrément SIAF ; certification DSA).

La plateforme PAC est conforme à la norme OAIS et s'inspire donc de ce modèle dans son architecture. Elle est aussi compatible avec le SEDA (Standard d'Echange de Données pour l'Archivage) pour l'archivage de données publiques et suit les recommandations de la norme Z 42-013.

Concrètement, la procédure d'archivage se déroule ainsi : le « service versant », après la collecte des documents auprès de la communauté de producteurs, les transmet par réseau à la plateforme. Une fois les documents validés, ils sont conservés dans le serveur de stockage, sous la responsabilité du service d'archives qui en assure l'administration. Tout document archivé est copié sur plusieurs supports (disques ou bandes) conservés dans des salles machines différentes. Une copie du document est sauvegardée sur site distant. Enfin, le serveur d'accès permet aux utilisateurs autorisés de consulter ces archives, et éventuellement d'obtenir une copie des documents archivés. À chaque objet archivé est attribué un identifiant unique et pérenne.

Les étapes et les acteurs du versement d'archives :



En pratique...

Le projet d'archivage doit être analysé et préparé en amont avec les personnes ressources de l'établissement ou du laboratoire (archiviste, documentalistes, services supports etc.). L'équipe de recherche doit contacter le CINES afin de présenter le projet, de décrire le type de données, les formats, la volumétrie etc. Le CINES procède ensuite à l'évaluation de la faisabilité du projet. En sciences humaines et sociales, la TGIR Huma-Num est l'entité compétente pour évaluer les demandes d'archivage et accompagner les équipes de recherche dans leur démarche. Le format des données doit être publié, libre et recevable par le CINES. L'outil FACILE offre la possibilité de tester la validité du format des fichiers conformément à la liste des formats archivables au CINES. Des tutoriels d'aide à la correction de PDF y sont également proposés.