**RÉPUBLIQUE FRANÇAISE**
Ministère de la culture et de la communication
Direction générale des patrimoines
Service interministériel des Archives de France

# Note d'information DGP/SIAF/2014/005 en date du 8 juillet 2014 relative à la journalisation des événements

Le directeur chargé des Archives de France

à

Mesdames et Messieurs les directeurs des services des Archives nationales
Mesdames et Messieurs les directeurs des services départementaux d'archives
sous couvert de Mesdames et Messieurs les préfets de région
et de Mesdames et Messieurs les préfets de département
Mesdames et Messieurs les responsables des services d'archives communales
sous couvert de Mesdames et Messieurs les maires
Mesdames et Messieurs les responsables des services d'archives régionales
sous couvert de Mesdames et Messieurs les présidents de région

La traçabilité est indispensable dans tout système d'information. Dans les plateformes d'archivage électronique, c'est une fonctionnalité centrale « métier » étant donnée la nature du système d'information (conservation à moyen, voire à long terme).

Toutefois, la norme OAIS pour la gestion, l'archivage et la préservation à long terme des documents numériques est une norme conceptuelle et de ce fait ne permet pas de répondre précisément à cette question. Quant à la norme Afnor NF Z 42-013 sur les spécifications pour la conception et l'exploitation des systèmes informatiques utilisés pour l'archivage électronique, si elle aborde bien la question de la traçabilité (journal des événements, journal du cycle de vie des données), à travers notamment son guide d'application GA Z 42-019, elle ne répond pas à l'ensemble des questions que l'on peut se poser.

Lors de la mise en œuvre d'un système, des questions restent donc en suspens :

- Quelles catégories d'informations souhaite-t-on tracer ?
- Comment exprimer ces informations dans le système et/ou comment les exprimer lorsqu'elles doivent sortir du système ?
- Quel cycle de vie pour ces informations ?

Dans ce contexte, le Service interministériel des Archives de France a réuni quelques acteurs publics du domaine de l'archivage électronique (programme VITAM, Centre Informatique

National de l'Enseignement Supérieur, Bibliothèque nationale de France, conseils généraux de la Gironde et de l'Aube) afin de pouvoir dresser un état de l'art. Les échanges ont permis de dégager des bonnes pratiques, qui font l'objet de la présente note d'information :

- identification des événements significatifs à journaliser ;
- liste des informations indispensables à enregistrer pour décrire un événement ;
- exemple d'implémentation en s'appuyant sur le bloc event du standard PREMIS (PREservation Metadata: Implementation Strategies) pour l'expression des métadonnées de préservation.

## Logs et journaux

Toutes les actions réalisées par une plate-forme d'archivage ne répondent pas aux mêmes objectifs et ne nécessitent pas toutes d'être enregistrées dans des journaux. Dans certains cas, la présence de ces informations dans des fichiers techniques peut s'avérer suffisante.

On peut distinguer deux grandes familles d'événements, du point de vue des usages associés :

- les événements qui permettent l'exploitation du système et servent d'abord à des fins de contrôle interne et que l'on appelle communément « logs » ;
- les événements qui affectent le cycle de vie de l'archive et qui permettent de délivrer, en cas de besoin, des attestations quant à la qualité (intégrité et authenticité) des archives conservées ; ces événements sont définis dans la norme Afnor NF Z 42-013 comme les opérations ayant lieu entre un déposant et le système d'archivage (création, modification des métadonnées, suppression et restitution). Cette notion sera désignée ici sous l'appellation de « journaux ».

## Événements à consigner dans les journaux

Les événements suivants sont considérés comme significatifs et il apparaît indispensable de les consigner dans les journaux :

- la validation du dépôt ;
- la migration de format ;
- la modification des métadonnées ;
- la restitution d'une archive ;
- la restauration d'une archive (pour tracer les opérations de correction suite à détection d'une anomalie) ;
- la suppression d'une archive ;

Les informations sur les autres opérations réalisées par la plate-forme d'archivage peuvent être conservées dans des fichiers logs, selon la politique de gestion des logs définie par l'organisme.

## Contenu des journaux

L'**action** se rapporte à une unité de gestion. Une unité de gestion, selon les cas, pourra correspondre à un paquet d'information archivé (AIP) ou à une collection d'informations archivées (AIC)[1]. L'unité de gestion doit nécessairement posséder un **identifiant** unique dans la plateforme d'archivage.

1 La norme OAIS distingue AIP et AIC, qu'elle définit ainsi :

L'action est réalisée par un ou des **agents**. L'agent à l'origine de l'action peut être une personne physique (par exemple l'administrateur de la plate-forme) ou un outil (par exemple le logiciel de conversion de format) ou encore une unité d'exécution portant des outils (par exemple un serveur hébergeant des outils de conversion de formats).

Seule la **date de fin d'action**, c'est-à-dire le **résultat de l'action**, est consignée dans le journal. Le résultat de l'action peut comporter trois valeurs : OK, Warning ou Error. Au moins dans les deux derniers cas, le résultat sera accompagné d'un **commentaire**. Les étapes qui ont conduit au résultat de l'action ne sont pas consignées dans le journal ; on peut en retrouver la trace dans les fichiers de logs : par exemple les contrôles intermédiaires réalisés en amont de la validation du transfert (contrôle de la taille du SIP, contrôle des formats, contrôle de l'unicité du transfert, etc.).

L'**empreinte de l'unité de gestion** après sa manipulation doit être consignée dans le journal à l'issue de l'action.

Une action, pour être réalisée, doit être encadrée par une règle (**contrat ou contexte de l'action**). Cette règle doit être référencée dans la plate-forme d'archivage.

Chaque événement peut donc être décrit de la manière suivante :

- action (type d'événement) ;
- agent à l'origine de l'action ;
-
- date de fin d'action ;
- identifiant de l'unité de gestion sur laquelle a porté l'action ;
- empreinte de l'unité de gestion résultant de l'action ;
- résultat de l'action (OK, Warning, Error) ;
- commentaire sur le résultat de l'action (en cas de résultat Warning ou Error) ;
- contrat ou contexte de l'action (identifiant de la règle qui autorise la réalisation de l'action).

### Forme des journaux

Les informations consignées dans le journal peuvent être exprimées dans différents formats : xml, csv par exemple.

Des équivalences entre les informations à enregistrer dans le journal et le bloc event de PREMIS sont proposées à titre d'exemple d'implémentation.

| Informations à enregistrer | PREMIS |
|---|---|
| Action | eventIdentifier (identifiant de l'événement)<br>– eventIdentifierType (type d'identifiant |

– 1/ Paquet d'informations archivé (Archival Information Package - AIP) : Paquet d'informations conservé dans un OAIS et constitué d'un Contenu d'information et de l'Information de pérennisation (PDI) associée.

– 2/ Collection d'informations archivée (Archival Information Collection - AIC) : Paquet d'informations archivé (AIP) dont le Contenu d'information est un agrégat d'autres AIP. La description fine de cette dernière notion se retrouve, en particulier, au paragraphe 4.2.2.7. de la norme.

| | |
|---|---|
| | de l'événement)<br>– eventIdentifierValue (valeur de l'identifiant de l'événement)<br><br>eventType (type d'événement) |
| Agent à l'origine de l'action | linkingAgentIdentifier (identifiant de l'agent lié) |
| Date de fin d'action | eventDateTime (date de l'événement) |
| Identifiant de l'unité de gestion sur laquelle a porté l'action | linkingObjectIdentifier (identifiant de l'objet lié) |
| Empreinte de l'unité de gestion résultant de l'action | linkingObjectIdentifier (identifiant de l'objet lié) |
| Résultat de l'action (OK, Warning, Error) | eventOutcome (résultat d'un événement) |
| Commentaire sur le résultat (Warning ou Error) | eventOutcomeDetail (détails sur le résultat d'un événement) |
| Contrat ou contexte de l'action | eventOutcomeDetailExtension (extension détaillant le résultat d'un événement) |

Les journaux doivent être chaînés pour garantir leur intégrité. Le guide d'application de la norme Afnor NF Z 42-013 propose un mode de chaînage qui est d'inscrire dans le journal n l'empreinte du journal n-1. D'autres modes opératoires peuvent être déclinés pour attester de la continuité des journaux et de leur intégrité, en fonction du niveau de service identifié. Par exemple, si les journaux sont archivés au rythme d'un par jour, l'horodatage est fourni par l'archivage et le chaînage par l'enregistrement journalier de ces journaux.

En cas de restitution des archives, les journaux correspondants devront être également restitués au producteur. Il conviendra d'être vigilant au mode opératoire de chainage des journaux de sorte que la restitution des journaux soit possible sans altérer leur continuité.

Le directeur, chargé des Archives de France

Hervé Lemoine