



Structuration des *Dissemination Information Packages* (DIP)

Date	Version
20/03/2018	3.0. (Release 6)

État du document

En projet Vérifié Validé

Maîtrise du document

Responsabilité	Nom	Entité	Date
Rédaction	NMO	Équipe Vitam	08/09/17
Vérification	Équipe	Équipe Vitam	
Validation		Équipe Vitam	20/03/2018

Suivi des modifications

Version	Date	Auteur	Modifications
0.1	14/11/17	NMO	Initialisation
1.0	28/11/17	MRE	Finalisation
2.0.	28/11/2017	MRE	Finalisation du document pour publication de la V1 fonctionnelle
3.0	20/03/2018	MRE	Finalisation du document pour publication de la V1 de production

Documents de référence

Document	Date	Remarques
Référentiel général de sécurité (RGS) – v. 2.0	10/06/2015	
ISO 14721:2012 – Systèmes de transfert des informations et données spatiales -- Système ouvert d'archivage d'information (SOAI) -- Modèle de référence	01/09/2012	
NF Z44-022 – MEDONA - Modélisation des données pour l'archivage	18/01/2014	
Standard d'échange de données pour l'archivage – SEDA – v. 2.0	31/12/2015	

Licence

La solution logicielle VITAM est publiée sous la licence CeCILL 2.1 ; la documentation associée (comprenant le présent document) est publiée sous Licence Ouverte V2.0.

Table des matières

1. Résumé.....	4
1.1 Présentation du programme Vitam.....	4
1.2 Présentation du document.....	5
2. Présentation du DIP.....	6
2.1. Définition.....	6
2.2 Contexte normatif.....	6
2.2.1 L’OAIS (Open Archival Information System).....	6
2.2.2 Le SEDA (Standard d’Échange de Données pour l’Archivage) et la norme MEDONA (Modèle d’Échange de DONnées pour l’Archivage).....	7
3. Constitution d’un <i>Dissemination Information Package</i> (DIP) par une plate-forme utilisant la solution logicielle Vitam.....	10
3.1. Les caractéristiques techniques du paquet à communiquer.....	10
3.2 Les modalités de lancement de l’opération de mise à disposition d’un DIP.....	10
3.3 Les modalités de récupération du DIP.....	11
3.4. Le bordereau.....	12
Annexe.....	13

1. Résumé

Jusqu'à présent, pour la gestion, la conservation, la préservation et la consultation des archives numériques, les acteurs du secteur public étatique ont utilisé des techniques d'archivage classiques, adaptées aux volumes limités dont la prise en charge leur était proposée. Cette situation évolue désormais rapidement et les acteurs du secteur public étatique doivent se mettre en capacité de traiter les volumes croissants d'archives numériques qui doivent être archivés, grâce à un saut technologique.

1.1 Présentation du programme Vitam

Les trois ministères (Europe et Affaires étrangères, Armées et Culture), combinant légalement mission d'archivage définitif et expertise archivistique associée, ont décidé d'unir leurs efforts, sous le pilotage de la Direction interministérielle du numérique et du système d'information et de communication de l'État (DINSIC), pour faire face à ces enjeux. Ils ont décidé de lancer un programme nommé Vitam (Valeurs Immatérielles Transmises aux Archives Pour Mémoire) qui couvre plus précisément les opérations suivantes :

- la conception, la réalisation et la maintenance mutualisées d'une solution logicielle d'archivage électronique de type back-office, permettant la prise en charge, le traitement, la conservation et l'accès aux volumes croissants d'archives (projet de solution logicielle Vitam) ;
- l'intégration par chacun des trois ministères porteurs du Programme de la solution logicielle dans sa plate-forme d'archivage. Ceci implique l'adaptation ou le remplacement des applications métiers existantes des services d'archives pour unifier la gestion et l'accès aux archives, la reprise des données archivées depuis le début des années 1980, la réalisation d'interfaces entre les applications productrices d'archives et la plate-forme d'archivage (projets SAPHIR au MEAE, ADAMANT au MC et ArchiPél au MA) ;
- le développement, par un maximum d'acteurs de la sphère publique, de politiques et de plates-formes d'archivage utilisant la solution logicielle (projet Ad-Essor).

La solution logicielle Vitam est développée en logiciel libre et recourt aux technologies innovantes du Big Data, seules à même de relever le défi de l'archivage du nombre d'objets numériques qui seront produits ces prochaines années par les administrations de l'État. Afin de s'assurer de la qualité du logiciel livré et de limiter les décalages calendaires de réalisation, le projet est mené selon une conduite de projet Agile. Cette méthode dite « itérative », « incrémentale » et « adaptative » opère par successions de cycles réguliers et fréquents de développements-tests-corrections-intégration. Elle associe les utilisateurs tout au long des développements en leur faisant tester les éléments logiciels produits et surtout en leur demandant un avis sur la qualité des résultats obtenus. Ces contrôles réguliers permettent d'éviter de mauvaises surprises lors de la livraison finale de la solution logicielle en corrigeant au fur et à mesure d'éventuels dysfonctionnements.

Le programme Vitam bénéficie du soutien du Commissariat général à l’investissement dans le cadre de l’action : « Transition numérique de l’État et modernisation de l’action publique » du Programme d’investissement d’avenir. Il a été lancé officiellement le 9 mars 2015, suite à la signature de deux conventions, la première entre les ministères porteurs et les services du Premier ministre, pilote du programme au travers de la DINSIC, et la seconde entre les services du Premier ministre et la Caisse des dépôts et consignations, relative à la gestion des crédits attribués au titre du Programme d’investissements d’avenir.

1.2 Présentation du document

Le présent document constitue une présentation des fonctionnalités associées au *Dissemination Information Package* dans la solution logicielle Vitam.

Le présent document décrit les fonctionnalités qui seront offertes par la solution logicielle Vitam au terme de la *release 6* (mars 2018). Il a vocation à être amendé, complété et enrichi au fur et à mesure de la réalisation de la solution logicielle Vitam et des retours et commentaires formulés par les ministères porteurs et les partenaires du programme.

2. Présentation du DIP

2.1. Définition

L'archivage a pour finalité la conservation de l'information en raison de sa valeur (administrative, juridique, historique, patrimoniale...), mais également sa mise à disposition. Ce processus permet au service d'archives de mettre à disposition de tiers (services producteurs, autres services d'archives, usagers sur place et à distance...) les archives qu'il conserve, si besoin après vérification des droits des demandeurs.

Dans l'univers numérique, peuvent être mis à disposition les documents eux-mêmes et leurs métadonnées.

Les services ainsi délivrés sont de plusieurs ordres :

- transmission d'ensembles de métadonnées à partir desquelles le service d'archives pourra élaborer des instruments de recherche pour faire connaître les ressources dont il dispose,
- communication des archives par consultation ou diffusion,
- réutilisation des archives sur les portails des services d'archives ou par des tiers,
- restitution d'archives au service producteur (sortie définitive en cas de réactivation d'un dossier ou de changement de prestataire d'archivage),
- transfert de responsabilité à un autre service d'archives, notamment à l'issue d'une phase d'archivage intermédiaire.

Tous ces services se traduisent par la constitution d'un lot d'archives à partir des éléments gérés par le service d'archives et la mise à disposition de ce lot.

Le présent document décrit la manière de préparer des paquets et de les mettre à disposition dans une plate-forme utilisant la solution logicielle Vitam.

2.2 Contexte normatif

La constitution d'un paquet pour mise à disposition par une plate-forme utilisant la solution logicielle Vitam doit être réalisée conformément aux deux normes suivantes.

2.2.1 L'OAIS (Open Archival Information System)

L'OAIS est un modèle conceptuel de référence décrivant dans les grandes lignes les fonctions, les responsabilités et l'organisation d'un système qui voudrait préserver sur le long terme de l'information, en particulier des données numériques.

Le modèle OAIS repose sur l'idée que l'information constitue des paquets, et que ces paquets ne sont pas les mêmes suivant qu'on est en train de produire l'information, de mettre en œuvre des opérations pour la conserver, ou de la communiquer à un utilisateur.

Il existe d'après cette norme trois sortes de paquets :

- **les paquets pris en charge par le système ou *Submission Information Packages***

(SIP) préparés par les producteurs à destination de la solution logicielle Vitam ;

- **les paquets gérés par le système ou *Archival Information Packages (AIP)*** transformés par la solution logicielle Vitam à partir du SIP dans une forme plus facile à conserver et à gérer dans le temps ;
- **les paquets communiqués par le système ou *Dissemination Information Package (DIP)*** transformés par la solution logicielle Vitam à partir d'un ou plusieurs AIP.

Un paquet d'archives destiné à être mis à disposition de tiers par une plate-forme utilisant la solution logicielle Vitam constitue un DIP au sens du modèle OAIS.

2.2.2 Le SEDA (Standard d'Échange de Données pour l'Archivage) et la norme MEDONA (Modèle d'Échange de DONnées pour l'Archivage)

La norme MEDONA et le standard SEDA constituent des modèles standards pour les transactions d'échanges entre les acteurs de l'archivage, et notamment entre le service d'archives et le demandeur d'une communication.

La norme AFNOR NF Z44-022, intitulée "Modèle d'Échange de DONnées pour l'Archivage" ou MEDONA a été publiée en 2014. Elle apporte un cadre normatif pour les différents échanges d'informations (données comme métadonnées) entre le service d'archives et ses partenaires comme le demandeur (toute personne physique ou morale qui souhaite consulter les informations conservées par le Service d'archives dans le respect des conditions légales, réglementaires ou contractuelles en vigueur) et la façon dont s'organisent ces échanges. Elle définit des diagrammes d'activités et des modèles de données selon le formalisme UML. Les échanges se traduisent par des messages formalisés par des schémas XML.

Le SEDA est le « Standard d'Échange de Données pour l'Archivage » relatif aux données d'archives publiques. Créé en 2006 par la Direction des Archives de France et la Direction Générale pour la Modernisation de L'État, il a été révisé en 2010 (v0.2), en 2012 (v1.0) et plus récemment en 2015 (v2.0) pour assurer sa compatibilité avec la norme MEDONA. Il comporte à la fois le modèle fonctionnel et un modèle de données pour les échanges de données d'archives entre acteurs de la sphère publique : description, acteurs, règles de gestion, référentiels métier et techniques.

Selon ce standard, les fonctions de mise à disposition couvertes par le DIP correspondent à deux cas d'utilisation : la communication et la restitution.

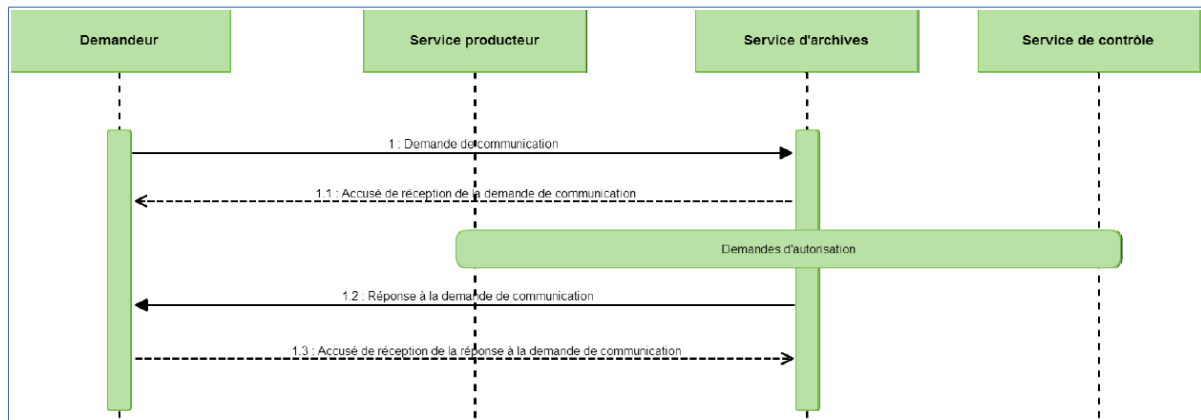
La demande de communication d'archives (ArchiveDeliveryRequest) est effectuée par un demandeur (le producteur de ces informations, ou un tiers) qui souhaite consulter des informations. La communication peut concerner tant les données elles-mêmes que leurs métadonnées.

À réception de la demande de communication, le service d'archives émet un accusé de réception en direction du demandeur.

Après instruction de la demande et si nécessaire une phase de demandes d'autorisations, une réponse (ArchiveDeliveryRequestReply) est envoyée par le service d'archives au demandeur. Cette réponse peut être négative (par exemple dans le cas où l'information demandée n'existe pas ou si le service de contrôle s'y oppose) ou être positive, auquel cas elle comprend les données demandées accompagnées de leurs métadonnées.

Une fois la réponse reçue, le demandeur émet un message d'accusé de réception.

Le schéma ci-dessous modélise les échanges associés à la requête et au transfert d'un ensemble d'archives à communiquer – un DIP au sens de la norme OAIS.



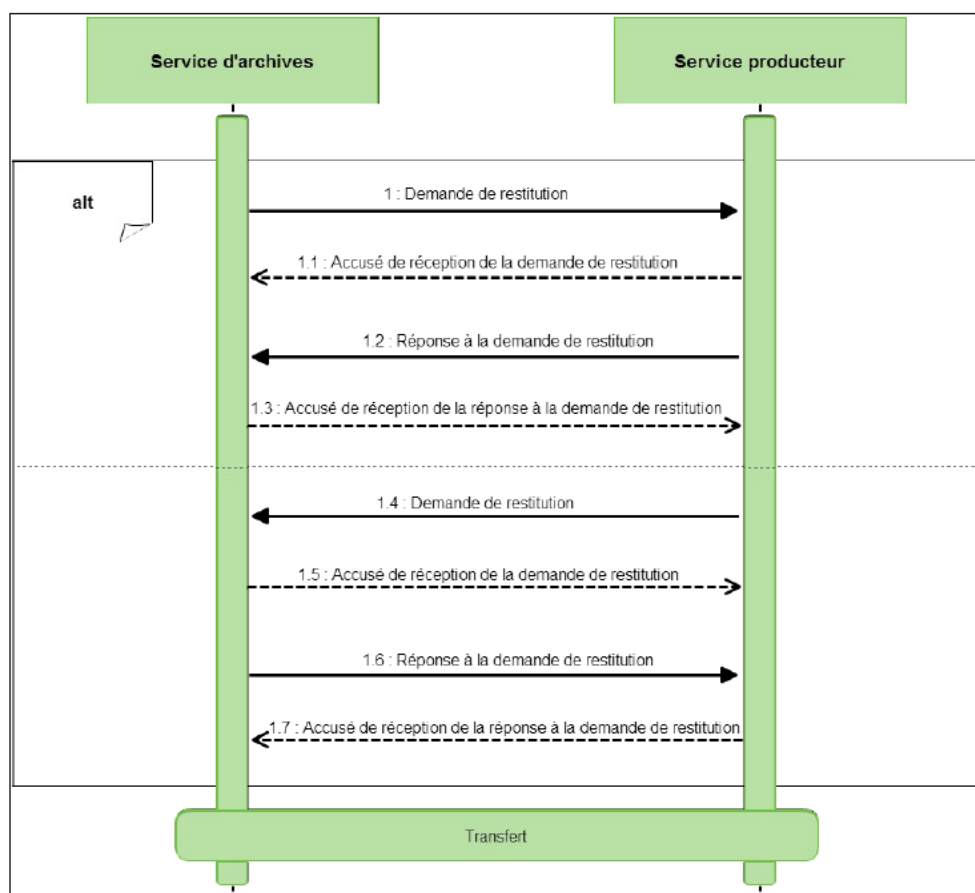
Point d'attention : la phase de demande d'autorisations est du ressort du front office et non du back office.

La transaction de restitution est découpée en deux séquences : une séquence de demande de restitution suivie d'une séquence de transfert.

La demande de restitution (ArchiveRestitutionRequest) peut être à l'initiative soit du service d'archives qui détient les informations à restituer, soit du service producteur.

La demande s'effectue en mentionnant les identifiants des données concernées (éventuellement accompagnées de leurs métadonnées). Le service qui reçoit la demande, en accuse la réception. S'ensuit l'envoi de sa réponse (acceptation ou refus de la demande de restitution - ArchiveRestitutionRequestReply) dont l'initiateur de la demande accuse la réception.

En cas d'acceptation de la demande de restitution, le transfert effectif des informations entre les acteurs suit le modèle normal d'un transfert.



Cependant, si le SEDA explique de manière générale à quoi ressemblent les processus de communication ou de restitution de données numériques, ce standard ne spécifie pas les règles de constitution du paquet à mettre à disposition.

C'est pourquoi le présent document vise à :

- apporter des précisions sur la constitution des DIP qui seront mis à disposition par une plate-forme utilisant la solution logicielle Vitam ;
- expliciter les choix faits par l'équipe projet Vitam quand le standard SEDA en laissait la responsabilité aux implémentations.

À défaut de précision apportée par le présent document, la documentation accompagnant le standard SEDA 2.0. et accessible sur le site internet du Service interministériel des Archives de France s'applique¹.

Dans le cadre de la *release 6* (mars 2018), l'équipe chargée de la réalisation de la solution logicielle Vitam a développé les fonctionnalités de base d'export d'un DIP : constitution du paquet d'objets-données (DataObjectPackage), mais sans encore implémenter les messages d'échanges prévus par le SEDA. Ces fonctionnalités de base seront enrichies dans les itérations ultérieures.

¹ Disponible à l'adresse suivante (lien vérifié le 14 novembre 2017) : <https://redirect.francearchives.fr/seda/>.

3. Constitution d'un *Dissemination Information Package* (DIP) par une plate-forme utilisant la solution logicielle Vitam

3.1. Les caractéristiques techniques du paquet à communiquer

D'après le SEDA, un DIP est constitué d'un bordereau et d'un ou plusieurs objets qui font l'objet de la demande de communication. Il peut être uniquement composé d'un bordereau, quand seules les métadonnées sont mises à disposition du demandeur.

Le DIP exporté par une plate-forme utilisant la solution logicielle Vitam répond aux caractéristiques suivantes.

Intitulé	Description	Niveau de recommandation
Structuration du DIP	Un DIP se compose : <ul style="list-style-type: none"> d'un bordereau nommé « manifest.xml » d'un répertoire nommé « Content » rassemblant le cas échéant les objets numériques à transférer au demandeur. En l'absence d'objet (par exemple si la demande ne porte que sur les métadonnées), la présence du répertoire « Content » n'est pas obligatoire.	Obligatoire
Type de format du DIP	ZIP.	Obligatoire
Composition du DIP	Les unités archivistiques du DIP doivent avoir le même service producteur (OriginatingAgency présente dans le ManagementMetadata du SIP).	Obligatoire
Taille du DIP	La taille maximale du DIP est de 10 000 unités archivistiques.	Obligatoire

3.2 Les modalités de lancement de l'opération de mise à disposition d'un DIP

Depuis l'IHM standard accompagnant la solution logicielle Vitam, à partir de la page de détail d'une unité archivistique, il est possible de demander à générer un DIP de l'unité archivistique uniquement ou de l'unité archivistique et de ses filles ou bien un DIP correspondant à l'opération d'entrée (ingest) initiale.

The screenshot shows a web interface titled "Génération du DIP". It contains three radio button options for selection:

- Unité archivistique
- Unité archivistique et sa descendance
- Ensemble de l'entrée

 Below these options is a purple button labeled "Exporter".

Via les API aux bornes de la solution logicielle Vitam, plusieurs critères sont utilisables : identifiant d'une opération d'entrée, unité archivistique précise, ensemble des unités archivistiques dépendant d'une unité archivistique précise.

Au lancement de l'opération, la solution logicielle Vitam génère un identifiant de l'opération. Le DIP constitué a pour nom cet identifiant d'opération.

L'opération est journalisée.

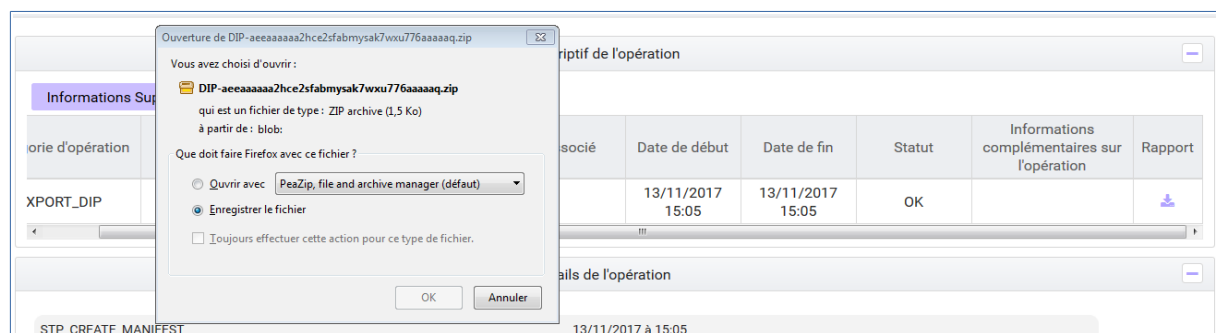
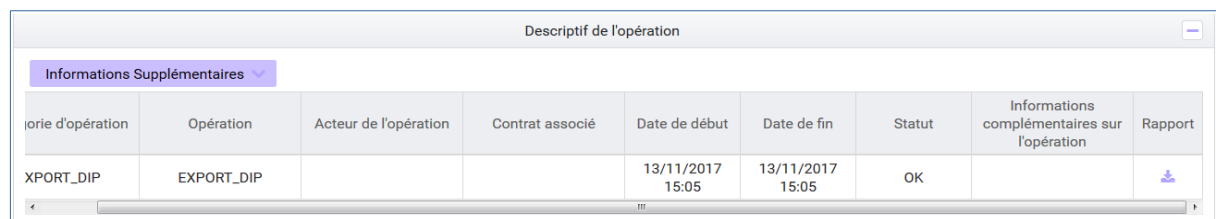
Points d'attention :

- dans le cas d'un DIP correspondant à une opération d'entrée, le paquet comprendra toutes les unités importées dans le système via cette opération d'entrée dans leur état au moment de la demande d'export du DIP : certaines unités archivistiques pourront avoir été modifiées depuis l'opération d'entrée initiale ;
- le DIP ne peut être généré qu'avec des unités archivistiques versées par un même service producteur.

3.3 Les modalités de récupération du DIP

Lorsque l'opération d'export du DIP est terminée, le DIP peut être récupéré :

- depuis l'IHM standard, à partir du journal des opérations en cliquant sur la ligne de l'opération d'export du DIP pour accéder à l'écran de détail, puis en affichant le champ « rapport » (via le sélecteur « Informations supplémentaires »),



- par API, en utilisant le service fourni par le endpoint `access-external/v1/dipexport`, au moyen de l'identifiant de l'opération d'export.

3.4. Le bordereau

À la racine du DIP se trouve le bordereau de mise à disposition qui décrit l'ensemble des métadonnées du paquet. Il est composé :

- d'une déclaration des objets binaires (DataObjectPackage > BinaryDataObject) ou des objets physiques (DataObjectPackage > PhysicalDataObject) ;
- d'une description des archives représentées par ces objets :
 - Métadonnées descriptives : DescriptiveMetadata ;
 - Métadonnées de gestion : ManagementMetadata ;
 - Arborescence intellectuelle : ArchiveUnit ;
- de la déclaration du service producteur (hérité de la balise OriginatingAgencyIdentifier présente dans le ManagementMetadata du SIP).

Annexe

Exemple de bordereau de mise à disposition

```
<ArchiveRestitutionRequest xmlns:xlink="http://www.w3.org/1999/xlink"
  xmlns:pr="info:lc/xmlns/premis-v2" xmlns="fr:gouv:culture:archivesdefrance:seda:v2.0"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="fr:gouv:culture:archivesdefrance:seda:v2.0 seda-2.0-main.xsd">
  <DataObjectPackage>
    <BinaryDataObject id="aeaaaaaaaaevotcgaavawak7x4dxlmiaaaba">
      <DataObjectGroupId>aebaaaaayevotcgaavawak7x4dxlmiaaabq</DataObjectGroupId>
      <DataObjectVersion>BinaryMaster_1</DataObjectVersion>
      <Uri>Content/aeaaaaaaaaevotcgaavawak7x4dxlmiaaaba</Uri>
      <MessageDigest algorithm="SHA-512"
        >59cca81fb1b51e80f27b644f959e7e0c2a9cabca6c5e437ada8b6f6517d845da99059c4fabb8
e59ac6dd791c7bd2c251149708b141f9a734cd144c5dfa0f4f70</MessageDigest>
      <Size>7405</Size>
      <FormatIdentification>
        <FormatLitteral>OpenDocument Text</FormatLitteral>
        <MimeType>application/vnd.oasis.opendocument.text</MimeType>
        <FormatId>fmt/291</FormatId>
      </FormatIdentification>
      <FileInfo>
        <Filename>Position_militaire.odt</Filename>
        <LastModified>2017-09-25T16:54:24.689+02:00</LastModified>
      </FileInfo>
    </BinaryDataObject>
    <BinaryDataObject id="aeaaaaaaaaevotcgaavawak7x4dxlmaaaada">
      <DataObjectGroupId>aebaaaaayevotcgaavawak7x4dxlmiaaaaq</DataObjectGroupId>
      <DataObjectVersion>BinaryMaster_1</DataObjectVersion>
      <Uri>Content/aeaaaaaaaaevotcgaavawak7x4dxlmaaaada</Uri>
      <MessageDigest algorithm="SHA-512"
        >59cca81fb1b51e80f27b644f959e7e0c2a9cabca6c5e437ada8b6f6517d845da99059c4fabb8
e59ac6dd791c7bd2c251149708b141f9a734cd144c5dfa0f4f70</MessageDigest>
      <Size>7405</Size>
      <FormatIdentification>
        <FormatLitteral>OpenDocument Text</FormatLitteral>
        <MimeType>application/vnd.oasis.opendocument.text</MimeType>
        <FormatId>fmt/291</FormatId>
      </FormatIdentification>
      <FileInfo>
        <Filename>ESS.odt</Filename>
        <LastModified>2017-09-25T16:54:38.491+02:00</LastModified>
      </FileInfo>
    </BinaryDataObject>
  </DataObjectPackage>
</ArchiveRestitutionRequest>
```

```
</BinaryDataObject>
<DescriptiveMetadata>
  <ArchiveUnit id="aeaqaaaaayevotcgaavawak7x4dxmqaaaea">
    <Management/>
    <Content>
      <DescriptionLevel>Item</DescriptionLevel>
      <Title xml:lang="fr">Etat signalétique et des services</Title>
      <FilePlanPosition>3</FilePlanPosition>
      <OriginatingAgencyArchiveUnitIdentifier>123456789</OriginatingAgencyArchiveUnitI
dentifier>
      <Addressee>
        <Corpname>DRH</Corpname>
      </Addressee>
      <AcquiredDate>1995-09-15</AcquiredDate>
      <ReceivedDate>1995-09-15</ReceivedDate>
    </Content>
    <DataObjectReference>
      <DataObjectGroupReferenceId>aebaaaaayevotcgaavawak7x4dxlmiaaaaq</DataObjectG
roupReferenceId>
    </DataObjectReference>
  </ArchiveUnit>
  <ArchiveUnit id="aeaqaaaaayevotcgaavawak7x4dxmqyaaaba">
    <Management/>
    <Content>
      <DescriptionLevel>Item</DescriptionLevel>
      <Title xml:lang="fr">Certificat de position militaire</Title>
      <FilePlanPosition>3</FilePlanPosition>
      <OriginatingAgencyArchiveUnitIdentifier>123456789</OriginatingAgencyArchiveUnitI
dentifier>
      <Addressee>
        <Corpname>DRH</Corpname>
      </Addressee>
      <AcquiredDate>1995-09-15</AcquiredDate>
      <ReceivedDate>1995-09-15</ReceivedDate>
    </Content>
    <DataObjectReference>
      <DataObjectGroupReferenceId>aebaaaaayevotcgaavawak7x4dxlmiaaabq</DataObject
GroupReferenceId>
    </DataObjectReference>
  </ArchiveUnit>
  <ArchiveUnit id="aeaqaaaaayevotcgaavawak7x4dxmqaaada">
    <Management/>
    <Content>
      <DescriptionLevel>File</DescriptionLevel>
      <Title xml:lang="fr">Situation militaire</Title>
```

```
<FilePlanPosition>3</FilePlanPosition>
  <OriginatingAgencyArchiveUnitIdentifier>123456789</OriginatingAgencyArchiveUnitI
dentifier>
  <StartDate>1995-09-15</StartDate>
  <EndDate>1996-03-12</EndDate>
</Content>
<ArchiveUnit id="aeaaaaaaaaevotcgaavawak7x5zkccqaaaq">
  <ArchiveUnitRefId>aeaqaayevotcgaavawak7x4dxmqaaaea</ArchiveUnitRefId>
</ArchiveUnit>
<ArchiveUnit id="aeaaaaaaaaevotcgaavawak7x5zkccqaaaba">
  <ArchiveUnitRefId>aeaqaayevotcgaavawak7x4dxmqaaaba</ArchiveUnitRefId>
</ArchiveUnit>
</ArchiveUnit>
</DescriptiveMetadata>
<ManagementMetadata>
  <OriginatingAgencyIdentifier>Vitam</OriginatingAgencyIdentifier>
</ManagementMetadata>
</DataObjectPackage>
</ArchiveRestitutionRequest>
```