

Archiving Mailboxes at the Archives nationales (France): overview and perspectives



Émeline Levasseur – emeline.levasseur@culture.gouv.fr
André Falut – andre.falut@culture.gouv.fr
Data migration and digital preservation managers



Department of data administration
Archives nationales (France)

1. Context



National Archives (France)

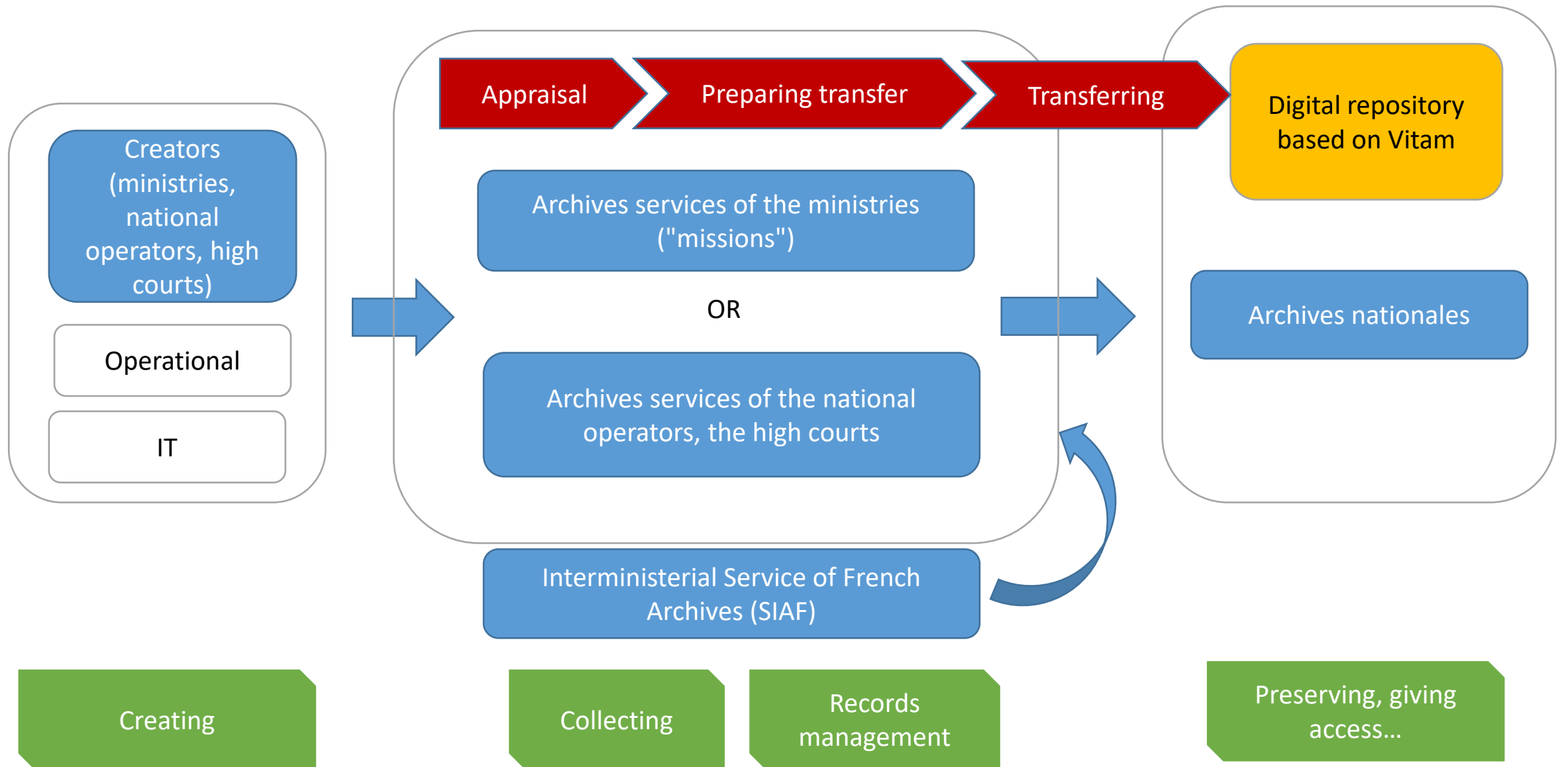
History: created during the French Revolution, on September 12th 1790

Missions: acquisition, arrangement, preserving, restoring, providing access, valorizing.



Scope: Historical public archives from the central administrations of State, archives of Paris notaries, private archives of national interest.

Organization of the archives network (national scale)



Digital archiving in French legal and normative context

Code du Patrimoine (2004, 2008):

- Defines archives as the documents, including data, whatever their date, their support, their form, etc., that are created or received by a person or an organization.
- Defines public archives.
- Sets up access status: records are open by default but restrictions may apply...

Open Archival Information System (OAIS)

NF Z 42-013 (2009, 2020): specifications - both organizational and technical - of a digital repository

Data Exchange Standard for Archiving – SEDA (2006, 2022): models the data exchanges (electronic and paper) between archival agencies and their partners, translated as messages in XML.

Technical context



Vitam: open-source software developed at its beginning with the sponsorship of 3 ministries (Culture, Defence, Foreign Affairs), now beyond (about 50 partners).

Goals: ensuring the ingestion, management and preservation of large amounts of digital records of administrative as well as archival value, while providing long-term access to these records.

The National Archives' digital repository is based on **Vitam (back-office)** and **interfaced with its own front-office** (in-house archives managing software).

In service since **November, 2018**.

Our digital repository is dedicated to born-digital archives. Digitized archives are excluded.

INTRANET

[Accueil](#) [Anomalie, que faire ?](#) [Actualités](#) [Archivage numérique](#)

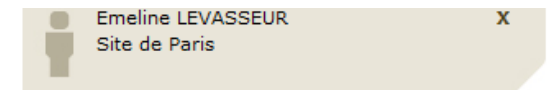
Gestion

- Fontainebleau
- Paris
- Pierrefitte-sur-Seine

Sélectionnez un module

Accès et diffusion

Sélectionnez un module



Rechercher

ARCHIVES NUMÉRIQUES

Accéder à l'ensemble des modules d'archivage numérique

Archives numériques



2. The National Archives' mailbox archiving strategy



Digital Archiving at the National Archives

Started in the early 80s, collecting **statistical** data which was preserved using a method called « **Constance** ».

The Constance method was based on the preservation of both **data** and its **representation information** (before OAIS!), as well as paper documentation (later digitized).

Data was stored on LTO tapes, representation information in in-house databases.

In the 2000s, digital archiving at the National Archives went on to encompass a wide variety of digital records: images, audio, film, files from office applications...

The Constance method was used until the launch of the digital repository (2018).



Encoding data in 1979 at the Archives nationales

See <https://www.archives-nationales.culture.gouv.fr/en/web/guest/archives-audiovisuelles-et-electroniques>

Mailboxes in our archival holdings

The first mailboxes entered our archival holdings in 2017.

Reminder: appraisal, transfer preparation and ingestion are conducted by mission archivists, except for private archives.

Mailboxes are generally **from ministries' offices**. Their digital records are also transferred.

In the digital repository: 40 PST files (representing 39 mailboxes) ingested.

Outside the digital repository, pending migration: 95 PST files, 7 mailboxes in Mbox.

=> Majority of PST files, as Outlook is widely used by the central administrations of State.

Mailbox archiving strategy

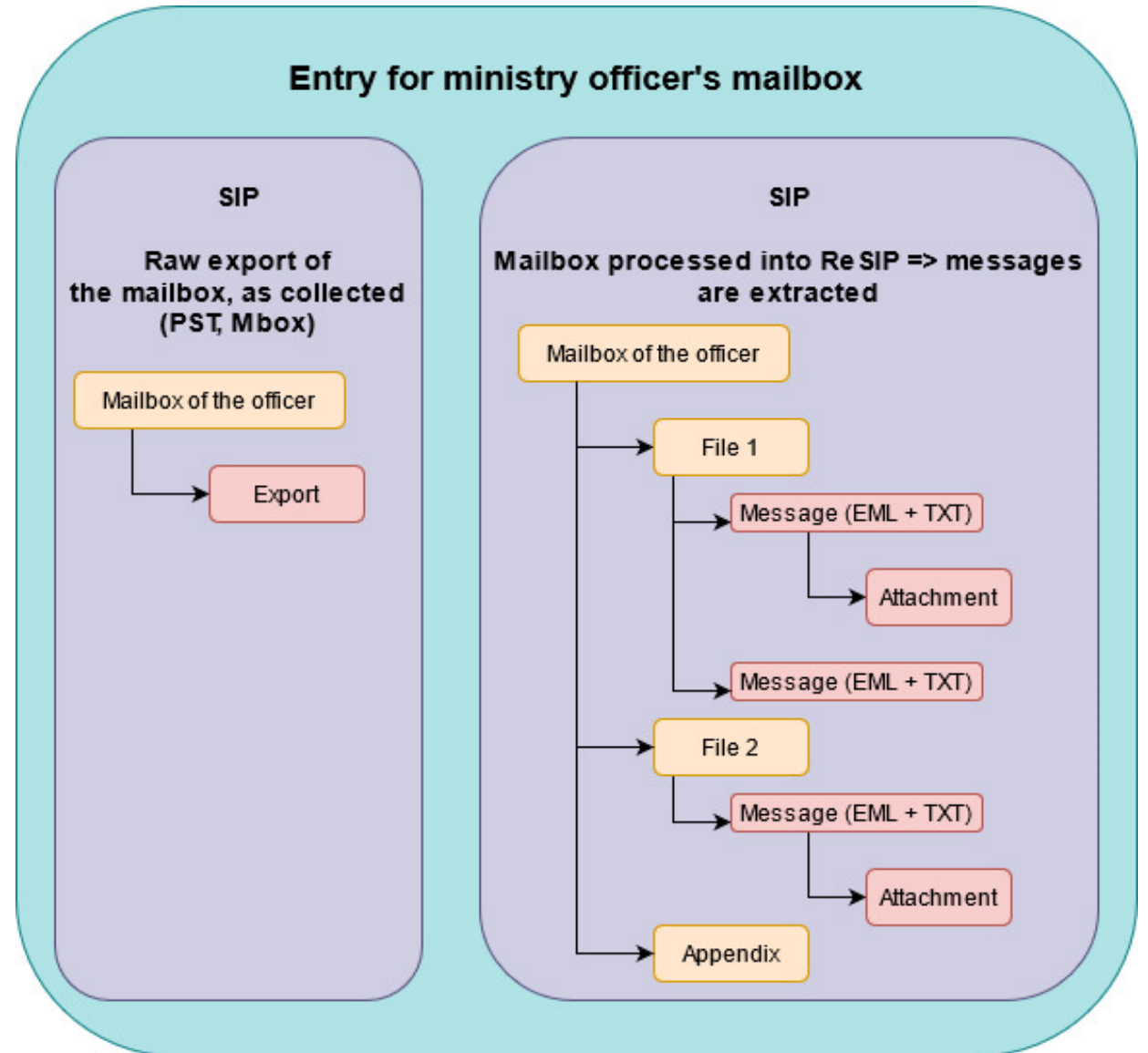
Goals:

- 1) **Ensuring preservation of the messages themselves**, their arrangement, their metadata and their attachments in their original format
- 2) **Making it possible to search and find messages** in the digital repository
- 3) **Enabling access** to the messages of a mailbox in a tree structure
- 4) **While leaving room for future options to represent mailboxes** and their messages outside the digital repository.

Mailbox archiving strategy

... In practice: 2 SIPs

1. The 1st contains the raw export (PST or Mbox) as received by the mission archivist
2. The 2nd contains:
 - The messages in EML format, rendered in their original tree structure
 - The body of each message in text format
 - A section of metadata extracted and mapped to SEDA entities
 - The extracted attachments linked to their messages.

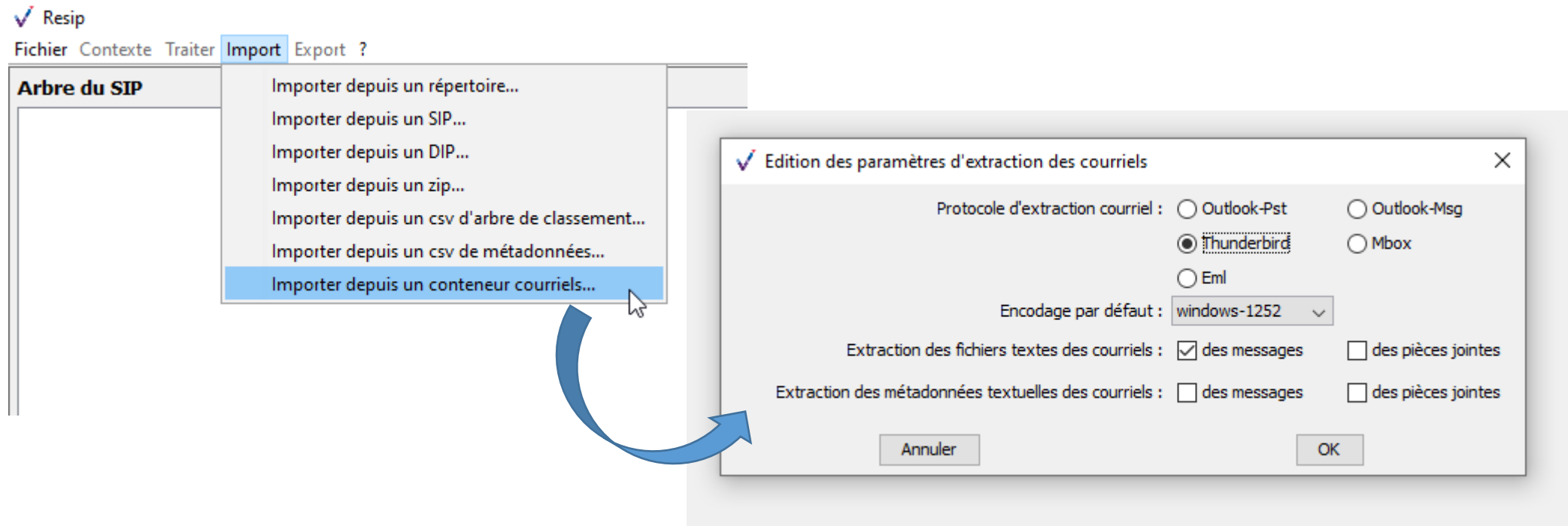


Resip

SIP preparation and mailbox extraction tool developed by the VITAM team.

Available on GitHub: <https://github.com/ProgrammeVitam/sedatools>

To import a mailbox:



Resip: mailbox importation results

Fichier Contexte Traiter Import Export ?

Extraction of the mailbox

+

A list of the messages

Arbre du STP (25663 archiveUnit/25567 dog/40618 bdo/0 pdo)

- Ensemble des messages électroniques et informations associées (contacts, rende...
- 3 Préparation des fonds (7/11597) - ID11
- 2 Etudes (6/776) - ID41105
- 5 Matériel de conservation (322/610) - ID43879
- Courrier entrant (2/312) - ID46035
- 4 Déménagement (11/7934) - ID47189
- Templates (1/1) - ID75879
- 1 chantier des fonds (5/2498) - ID75884
- 7 Politique scientifique (4/7) - ID84865
- 6 débats (loi_cité_MHF) (1187/1917) - ID84891
- messages.csv (0/0) - ID91857

Unité d'archive - ID10 ...

Descriptif :

Niveau de description :- RecordGrp

Titre +lang :- Ensemble des messages électroniques et informations ass

Date de début :- 5 mars 2007 09:42

Date de fin :- 28 novembre 2011 13:26

Recharger unité d'archive Sauver unité d'archive

Unité d'archive sans groupe d'objets

Ajouter un groupe d'objets

Recharger groupe d'objets

+ (direct subAU/total subAU) - xmlID

Ouvrir dossier AU/OG

Visualisation du contexte d'import

resip: extraction depuis un conteneur courriel sur disque
en [C:\Users\emeline.levasseur\Documents\01_DAD\messenger\atelier_pele_mel_20230214\local_folders-1]
encodé selon le format [thunderbird]
génération de fichier en forme texte des messages: non
extraction des métadonnées texte des messages: non
génération de fichier en forme texte des pièces jointes: non
extraction des métadonnées texte des pièces jointes: non
résultat: PT2M20.025S writing 95 folders, 13710 messages with 233 messages embedded for a total size of 1,5 GBen 2M20.063S

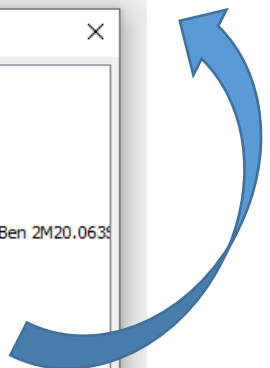
Résumé :

- ArchiveTransfer
- 23745 ArchiveUnit(s)
- 21882 DataObjectGroup(s)
- 21882 BinaryDataObject(s) (2,6 GB)
- 0 PhysicalDataObject(s)

encodé selon un modèle V2 de la structure
chargé en 2M26.698S

Vérification de l'absence de cycle et normalisation Vitam effectuées
Pas de modification

Fermer



Arbre du SIP (25663 archiveUnit/ 25567 dog/ 40618 bdo/ 0 pdo)

- Ensemble des messages électroniques et informations associées (contacts, rer)
 - 3 Préparation des fonds (7/11597) - ID11
 - 1 Programmation9f982eba (319/610) - ID12
 - [Fwd: Re: Chantiers des Fonds - suite réunion du 19-1 0] (0/0) - ID13
 - DataObjectGroup ID14
 - Re: préprogrammation 2008 (1/1) - ID17
 - préprogrammation (1/1) - ID24
 - Re: Chantiers des Fonds - suite réunion du 19-10 (1/1) - ID31
 - Microfilmage (2/2) - ID38
 - [Fwd: Re: Chantiers des Fonds - suite réunion du 19-10] (1/1) - ID39
 - Fiche d'évaluation (1/1) - ID55
 - Re: Chantiers des Fonds - suite réunion du 19-10 (4/4) - ID62
 - Envoi d'un message : Préprogrammation2008- F7.doc (1/1) - ID78
 - Envoi d'un message : Préprogr2008 Identité judiciaire-3.doc (1/1) - ID79
 - [Fwd: Re: Chantiers des Fonds - suite réunion du 19-10] (4/4) - ID80
 - Re: préprogrammation 2008 (0/0) - ID108
 - préprogrammation 2008 (0/0) - ID112
 - [Fwd: Re: Chantiers des Fonds - suite réunion du 19-10] (0/0) - ID113
 - [Fwd: Re: Chantiers des Fonds - suite réunion du 19-10] (0/0) - ID114
 - IJ et plaques (0/0) - ID124
 - Projet numérisation (1/1) - ID128
 - [Fwd: réunion du 19 octobre] (0/0) - ID135
 - Planification (1/1) - ID139
 - Planification Archives nationales (1/1) - ID146
 - quelques points en suspens (0/0) - ID153
 - programmation et budget 2008 (3/3) - ID157
 - planification (2/2) - ID170
 - programmation 2008 (0/0) - ID180
 - Re: pour avis (0/0) - ID184
 - Re: IJ et plaques (0/0) - ID188
 - album du parc (0/0) - ID192
 - Re: préprogrammation 2008 (0/0) - ID196
 - Re: Opération numérisation Photos Expo Universelle 1867 (0/0) - ID200
 - Re: conditionnement (0/0) - ID204
 - Re: réunion conditionnement (0/0) - ID208
 - réunion conditionnement (0/0) - ID212
 - Re: réunion conditionnement (0/0) - ID216

☑ + (direct subAU/total subAU) - xmlID

Ouvrir dossier AU/OG

Unité d'archive - ID13 ...

Descriptif :

Niveau de description : - Item

Titre [+lang] : - [Fwd: Re: Chantiers des Fonds - suite réunion du 19-1 0]

ID-SI-prod : + <472EDCA8.2040401@culture.gouv.fr>

Rédacteur : - francoise peyre, francoise.peyre@culture.gouv.fr

Destinataire : - christele noulet, christele.noulet@culture.gouv.fr

Date d'envoi : - 5 novembre 2007 09:04

Date de réception : - 5 novembre 2007 09:04

Recharger unité d'archive

Sauver unité d'archive

Groupe d'objets - ID14 ...

- > Numérique - ID15 : - BinaryMaster_1, -472EDCA8.2040401-culture.gouv.f.eml, mes
- > Numérique - ID16 : - TextContent_1, -472EDCA8.2040401-culture.gouv.f.txt, text

Recharger groupe d'objets

Sauver groupe d'objets

Arbre du SIP (25663 archiveUnit/25567 dog/40618 bdo/0 pdo)

Ensemble des messages électroniques et informations associées (contacts, rer

- 3 Préparation des fonds (7/11597) - ID11
 - 1 Programmation9f982eba (319/610) - ID12
 - [Fwd: Re: Chantiers des Fonds - suite réunion du 19-1 0] (0/0) - ID13
 - DataObjectGroup ID14
 - Re: préprogrammation 2008 (1/1) - ID17
 - préprogrammation (1/1) - ID24
 - Re: Chantiers des Fonds - suite réunion du 19-10 (1/1) - ID31
 - Microfilmage (2/2) - ID38
 - [Fwd: Re: Chantiers des Fonds - suite réunion du 19-10] (1/1) - ID40
 - Fiche d'évaluation (1/1) - ID55
 - Re: Chantiers des Fonds - suite réunion du 19-10 (4/4) - ID62
 - Envoi d'un message : Préprogrammation2008- F7.doc (1/1) - ID78
 - Envoi d'un message : Préprogr2008 Identité judiciaire-3.doc (1/1) - ID85
 - [Fwd: Re: Chantiers des Fonds - suite réunion du 19-10] (4/4) - ID92
 - Re: préprogrammation 2008 (0/0) - ID108
 - préprogrammation 2008 (0/0) - ID112
 - [Fwd: Re: Chantiers des Fonds - suite réunion du 19-10] (0/0) - ID118
 - [Fwd: Re: Chantiers des Fonds - suite réunion du 19-10] (0/0) - ID124
 - IJ et plaques (0/0) - ID124
 - Projet numérisation (1/1) - ID128
 - [Fwd: réunion du 19 octobre] (0/0) - ID135
 - Planification (1/1) - ID139
 - Planification Archives nationales (1/1) - ID146
 - quelques points en suspens (0/0) - ID153
 - programmation et budget 2008 (3/3) - ID157
 - planification (2/2) - ID170
 - programmation 2008 (0/0) - ID180
 - Re: pour avis (0/0) - ID184
 - Re: IJ et plaques (0/0) - ID188
 - album du parc (0/0) - ID192
 - Re: préprogrammation 2008 (0/0) - ID196
 - Re: Opération numérisation Photos Expo Universelle 1867 (0/0) - ID200
 - Re: conditionnement (0/0) - ID204
 - Re: réunion conditionnement (0/0) - ID208
 - réunion conditionnement (0/0) - ID212
 - Re: réunion conditionnement (0/0) - ID216

 + (direct subAU/total subAU) - xmlID


Ouvrir dossier AU/OG

Unité d'archive - ID13

```

1 <Content>
2 <DescriptionLevel>Item</DescriptionLevel>
3 <Title>[Fwd: Re: Chantiers des Fonds - suite réunion du 19-1 0]</Title>
4 <OriginatingSystemId>&lt;472EDCA8.2040401@culture.gouv.fr></OriginatingSystemId>
5 <Writer>
6 <FullName>francoise peyre</FullName>
7 <Identifieur>francoise.peyre@culture.gouv.fr</Identifieur>
8 </Writer>
9 <Addressee>
10 <FullName>christele noulet</FullName>
11 <Identifieur>christele.noulet@culture.gouv.fr</Identifieur>
12 </Addressee>
13 <SentDate>2007-11-05T09:04:40</SentDate>
14 <ReceivedDate>2007-11-05T09:04:42</ReceivedDate>
15 </Content>

```

Editer

Ajouter...

[A]ArchiveUnitProfile

Groupe d'objets

ID15-BinaryMaster_1
ID16-TextContent_1

Ouvrir l'objet

Changer l'objet

Détails

```

<DataObjectVersion>BinaryMaster_1</DataObjectVersion>
<Uri>content/ID15.eml</Uri>
<MessageDigest
algorithm="SHA-512">d447306fde5601bd20643a6b66e71cfc7:
<Size>6091</Size>
<FormatIdentification>
<FormatLitteral>MIME Email</FormatLitteral>
<MimeType>message/rfc822</MimeType>
<FormatId>fmt/950</FormatId>
</FormatIdentification>
<FileInfo>
<Filename>-472EDCA8.2040401-culture.gouv.f.eml</Filena
<LastModified>2007-11-05T12:25:15.815400</LastModified>

```

Editer le DataObject

Arbre du SIP (25663 archiveUnit/25567 dog/40618 bdo/0 pdo)

- Ensemble des messages électroniques et informations associées (contacts, rer)
 - 3 Préparation des fonds (7/11597) - ID11
 - 1 Programmation9f982eba (319/610) - ID12
 - [Fwd: Re: Chantiers des Fonds - suite réunion du 19-10] (0/0) - ID16
 - Re: préprogrammation 2008 (1/1) - ID17**
 - Préprogrammation2008.doc (0/0) - ID18
 - DataObjectGroup ID21
- préprogrammation (1/1) - ID24
- Re: Chantiers des Fonds - suite réunion du 19-10 (1/1) - ID31
- Microfilmage (2/2) - ID38
- [Fwd: Re: Chantiers des Fonds - suite réunion du 19-10] (1/1) - ID42
- Fiche d'évaluation (1/1) - ID55
- Re: Chantiers des Fonds - suite réunion du 19-10 (4/4) - ID62
- Envoi d'un message : Préprogrammation2008- F7.doc (1/1) - ID78
- Envoi d'un message : Préprogr2008 Identité judiciaire-3.doc (1/1)
- [Fwd: Re: Chantiers des Fonds - suite réunion du 19-10] (4/4) - ID84
- Re: préprogrammation 2008 (0/0) - ID108
- préprogrammation 2008 (0/0) - ID112
- [Fwd: Re: Chantiers des Fonds - suite réunion du 19-10] (0/0) - ID116
- [Fwd: Re: Chantiers des Fonds - suite réunion du 19-10] (0/0) - ID117
- IJ et plaques (0/0) - ID124
- Projet numérisation (1/1) - ID128
- [Fwd: réunion du 19 octobre] (0/0) - ID135
- Planification (1/1) - ID139
- Planification Archives nationales (1/1) - ID146
- quelques points en suspens (0/0) - ID153
- programmation et budget 2008 (3/3) - ID157
- planification (2/2) - ID170
- programmation 2008 (0/0) - ID180
- Re: pour avis (0/0) - ID184
- Re: IJ et plaques (0/0) - ID188
- album du parc (0/0) - ID192
- Re: préprogrammation 2008 (0/0) - ID196
- Re: Opération numérisation Photos Expo Universelle 1867 (0/0) - ID200
- Re: conditionnement (0/0) - ID204
- Re: réunion conditionnement (0/0) - ID208
- réunion conditionnement (0/0) - ID212

 + (direct subAU/total subAU) - xmlID

Ouvrir dossier AU/OG

Unité d'archive - ID17

```

1 <Content>
2 <DescriptionLevel>Item</DescriptionLevel>
3 <Title>Re: préprogrammation 2008</Title>
4 <OriginatingSystemId>&lt;47283C4B.1070200@culture.gouv.fr></OriginatingSystemId>
5 <Writer>
6 <FullName>marie-paule schmitt</FullName>
7 <Identifieur>marie-paule.schmitt@culture.gouv.fr</Identifieur>
8 </Writer>
9 <Adressee>
10 <FullName>christele noulet</FullName>
11 <Identifieur>christele.noulet@culture.gouv.fr</Identifieur>
12 </Adressee>
13 <Recipient>
14 <FullName>christine petillat</FullName>
15 <Identifieur>christine.petillat@culture.gouv.fr</Identifieur>
16 </Recipient>
17 <SentDate>2007-10-31T08:26:51Z</SentDate>
18 <ReceivedDate>2007-10-31T08:26:43Z</ReceivedDate>
19 <OriginatingSystemIdReplyTo>&lt;47220227.2060101@culture.gouv.fr></OriginatingSystemIdReplyTo>
20 </Content>

```

Editer

Ajouter...

[A]ArchiveUnitProfile

Groupe d'objets

ID22-BinaryMaster_1
ID23-TextContent_1

Ouvrir l'objet

Changer l'objet

Détails

```

<DataObjectVersion>BinaryMaster_1</DataObjectVersion>
<Uri>content/ID22.eml</Uri>
<MessageDigest
algorithm="SHA-512">a0639dc1cbabb2828f736c09802eeba0c<
<Size>44089</Size>
<FormatIdentification>
<FormatLitteral>MIME Email</FormatLitteral>
<MimeType>message/rfc822</MimeType>
<FormatId>fmt/950</FormatId>
</FormatIdentification>
<FileInfo>
<Filename>-47283C4B.1070200-culture.gouv.f.eml</Filename>
<LastModified>2007-10-31T12:25:16.822281</LastModified>

```

Editer le DataObject

Arbre du SIP (25663 archiveUnit/25567 dog/40618 bdo/0 pdo)

- Ensemble des messages électroniques et informations associées (contacts, rer ^
 - 3 Préparation des fonds (7/11597) - ID11
 - 1 Programmation9f982eba (319/610) - ID12
 - [Fwd: Re: Chantiers des Fonds - suite réunion du 19-10] (0/0) - ID109
 - Re: préprogrammation 2008 (1/1) - ID17
 - Préprogrammation2008.doc (0/0) - ID18
 - DataObjectGroup ID21
 - préprogrammation (1/1) - ID24
 - Re: Chantiers des Fonds - suite réunion du 19-10 (1/1) - ID31
 - Microfilmage (2/2) - ID38
 - [Fwd: Re: Chantiers des Fonds - suite réunion du 19-10] (1/1) - ID40
 - Fiche d'évaluation (1/1) - ID55
 - Re: Chantiers des Fonds - suite réunion du 19-10 (4/4) - ID62
 - Envoi d'un message : Préprogrammation2008- F7.doc (1/1) - ID78
 - Envoi d'un message : Préprogr2008 Identité judiciaire-3.doc (1/1)
 - [Fwd: Re: Chantiers des Fonds - suite réunion du 19-10] (4/4) - ID81
 - Re: préprogrammation 2008 (0/0) - ID108
 - préprogrammation 2008 (0/0) - ID112
 - [Fwd: Re: Chantiers des Fonds - suite réunion du 19-10] (0/0) - ID113
 - [Fwd: Re: Chantiers des Fonds - suite réunion du 19-10] (0/0) - ID114
 - IJ et plaques (0/0) - ID124
 - Projet numérisation (1/1) - ID128
 - [Fwd: réunion du 19 octobre] (0/0) - ID135
 - Planification (1/1) - ID139
 - Planification Archives nationales (1/1) - ID146
 - quelques points en suspens (0/0) - ID153
 - programmation et budget 2008 (3/3) - ID157
 - planification (2/2) - ID170
 - programmation 2008 (0/0) - ID180
 - Re: pour avis (0/0) - ID184
 - Re: IJ et plaques (0/0) - ID188
 - album du parc (0/0) - ID192
 - Re: préprogrammation 2008 (0/0) - ID196
 - Re: Opération numérisation Photos Expo Universelle 1867 (0/0) - ID200
 - Re: conditionnement (0/0) - ID204
 - Re: réunion conditionnement (0/0) - ID208
 - réunion conditionnement (0/0) - ID212

 + (direct subAU/total subAU) - xmlID


Ouvrir dossier AU/OG

Unité d'archive - ID18

```

1 <Content>
2   <DescriptionLevel>Item</DescriptionLevel>
3   <Title>Préprogrammation2008.doc</Title>
4   <Description>Document "Préprogrammation2008.doc joint</Description>
5 </Content>

```

Editer

Ajouter...

[A]ArchiveUnitProfile

Groupe d'objets

ID20-BinaryMaster_1

Ouvrir l'objet

Changer l'objet

Détails

```

<DataObjectVersion>BinaryMaster_1</DataObjectVersion>
<Uri>content/ID20.doc</Uri>
<MessageDigest
algorithm="SHA-512">0612090015f6461cd7df663a44c27ec83c
<Size>29696</Size>
<FormatIdentification>
  <FormatLiteral>Microsoft Word Document</FormatLiteral>
  <MimeType>application/msword</MimeType>
  <FormatId>fmt/40</FormatId>
</FormatIdentification>
<FileInfo>
  <Filename>Préprogrammation2008.doc</Filename>
  <LastModified>2001-05-21T12:25:16.82281</LastModified>

```

Editer le DataObject

Resip: appraisal tools

Statistiques

Par catégories de format

Catégorie	Nombre	Taille min
Tous formats	40618	0
Messagerie (mbox,pst,eml...)	15158	568
Texte brut	15059	1
Texte (doc,docx,odt...)	2879	0
Pdf	2068	4 573
Image (jpg,jpg2000,tiff...)	1945	0
Autres...	1835	0
Tableur (csv,xls,xlsx,ods...)	1421	783
Compressé (zip,tar...)	139	5 390
Présentation (ppt,pptx,odp...)	62	14 268
HTML	48	992
Non connu	4	111
Base de données (access,filemaker...)	0	-
Chiffré	0	-
Dessin (svg,odg,autocad...)	0	-

Attention, votre contexte de travail contient des fichiers vides ce qui n'est pas acceptable. Veuillez à les supprimer pour construire un SIP valide.

Chercher des doublons

Critères de sélection des doublons de groupes d'objets

- pour les objets binaires : Hachage Nom

- pour les objets physiques : Métadonnées

997 lots de DOG doublons/5374 AU
Aucun lot sélectionné

Index	Nb AU	Nb DOG	Noms	Tailles	Formats	MimeTypes
879	325	325	FRAN-signature-image-...	29 294	fmt/44	image/jpeg
244	191	191	clotilde-leforestier.vcf	257	fmt/395	text/vcard
312	186	186	FRAN-2011-COM-signat...	38 170	fmt/44	image/jpeg
31	170	170	clotilde-leforestier.vcf	342	fmt/395	text/vcard
53	153	153	FRAN-signature-image-...	21 073	fmt/44	image/jpeg
721	152	152	FRAN-signature-image-...	34 780	fmt/44	image/jpeg
449	132	132	clotilde-leforestier.vcf	357	fmt/395	text/vcard
259	119	119	clotilde-leforestier.vcf	410	fmt/395	text/vcard
148	108	108	anne-lama.vcf	369	fmt/395	text/vcard
514	105	105	FRAN-signature-image-...	78 238	fmt/44	image/jpeg
878	105	105	caroline-bongard.vcf	396	fmt/395	text/vcard
599	89	89	eleonore-alquier.vcf	288	fmt/395	text/vcard
70	78	78	caroline-bongard.vcf	601	fmt/395	text/vcard
471	78	78	caroline-bongard.vcf	544	fmt/395	text/vcard
400	76	76	anne-emilie-andres.vcf	235	fmt/395	text/vcard
843	70	70	caroline-bongard.vcf	379	fmt/395	text/vcard
923	65	65	FRAN-signature-image-f...	45 232	fmt/44	image/jpeg
433	61	61	noname	26 178	fmt/44	image/jpeg

Fusionner la ligne Fusionner tous les doublons

Chercher des unités d'archives

confidentiel

RegExp Respecter la casse Métadonnées ID 2/19 trouvés

Sans descendant : Unité d'archive Objet numérique ou physique

Resip: extraction settings

✓ Edition des paramètres par défaut ✕

Métadonnées globales | Métadonnées globales étendues | Export | Import | Traitement/Interface

Import des messageries

Encodage par défaut des messageries : windows-1252

Extraction des fichiers textes des courriels : des messages des pièces jointes

Extraction des métadonnées textuelles des courriels : des messages des pièces jointes

Import des hiérarchies sur disque

Fichiers exclus des imports :

```
Thumbs.db
pagefile.sys
.*\vcf
.*\emz
.*\dat
```

ignorer les liens symboliques et raccourcis

Import/export des csv

Encodage des csv : UTF-8

Séparateur : ;

Annuler OK

Resip: archival description and arrangement

Unité d'archive - ID10 ...

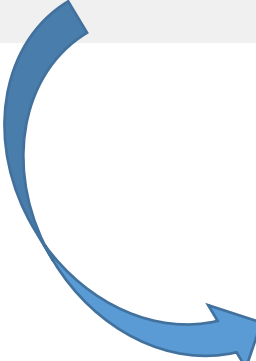
▼ Descriptif : - ...

Niveau de description :- RecordGrp ▼

Titre +lang :-+ Messagerie de Christelle Noulet, archiviste aux Archives nationales

Date de début :- 5 mars 2007 ... 09:42 ▼

Date de fin :- 28 novembre 2011 ... 13:26 ▼



✓ Resip

Fichier Contexte Traiter Import Export ?

Arbre du SIP (25663 archiveUnit/25567 dog/40618 bdo/0 pdo)

- Message de Christelle Noulet, archiviste aux Archives nationales (10/25662) - I
 - 3 Préparation des fonds (7/11597) - ID11
 - 1 Programmation9f982eba (319/610) - ID12
 - 5 numérisation (13/3076) - ID2166
 - 4 microfilmage (15/544) - ID13199
 - 3 conditionnement (37/6313) - ID15136
 - 2 Suivi des opérations (346/1044) - ID37608
 - Opération registres : V1 du CCTP (2/2) - ID41088
 - Re: Opération registres : V1 du CCTP (1/1) - ID41098
 - 2 Etudes (6/776) - ID41105
 - 5 Matériel de conservation (322/610) - ID43879
 - Courrier entrant (2/312) - ID46035
 - 4 Déménagement (11/7934) - ID47189
 - Templates (1/1) - ID75879
 - 1 chantier des fonds (5/2498) - ID75884
 - 7 Politique scientifique (4/7) - ID84865
 - 6 débats (loi, cité, MHF) (1187/1917) - ID84891
 - messages.csv (0/0) - ID91857

Unité d'archive - ID10 ...

▼ Descriptif : - ...

Niveau de description :- RecordGrp ▼

Titre +lang :-+ Messagerie de Christelle Noulet, archiviste aux Archives n:

Date de début :- 5 mars 2007 ... 09:42 ▼

Date de fin :- 28 novembre 2011 ... 13:26 ▼

Arbre du SIP (25663 archiveUnit/25567 dog/40618 bdo/0 pdo)

- Messagerie de Christelle Noulet, archiviste aux Archives nationales (10/25662) - I
 - 3 Préparation des fonds (7/11597) - ID11
 - 1 Programmation9f982eba (319/610) - ID12
 - 5 numérisation (13/3076) - ID2166
 - 4 microfilmage (15/544) - ID13199
 - 3 conditionnement (37/6313) - ID15136
 - 2 Suivi des opérations (346/1044) - ID37608
 - Opération registres : V1 du CCTP (2/2) - ID41088
 - Re: Opération registres : V1 du CCTP (1/1) - ID41098
 - 2 Etudes (6/776) - ID41105
 - 5 Matériel de conservation (322/610) - ID43879
 - Courrier entrant (2/312) - ID46035
 - 4 Déménagement (11/7934) - ID47189
 - Templates (1/1) - ID75879
 - 1 chantier des fonds (5/2498) - ID75884
 - 7 Politique scientifique (4/7) - ID84865
 - 6 débats (loi, cité, MHF) (1187/1917) - ID84891
 - messages.csv (0/0) - ID91857

Unité d'archive - ID10 ...

▼ Descriptif :

Niveau de description : ...

Titre +lang : ...

Date de début : ...

Date de fin : ...

- Contenu textuel
- Couverture
- Date d'enregistrement
- Date d'envoi
- Date de création
- Date de numérisation
- Date de réception
- Date de transaction
- Description
- Destinataire
- Destinataire-copie
- Emetteur
- Etat
- Evènement
- Expéditeur
- GPS
- Historique conservation
- ID-Re-SI-prod
- ID-SI-prod
- ID-archivreur
- ID-classement-prod
- ID-interne
- ID-papier
- ID-producteur
- ID-versant
- Langue
- Langue de description
- Mot-clé
- Objet associé
- Rédacteur
- Service producteur
- Service versant

Unité d'archive - ID10 ...

▼ Descriptif :

Niveau de description :- RecordGrp

Titre +lang :-+ Messagerie de Christelle Noulet, archiviste aux Archives n

Description +lang :-+ Description of the mailbox

Date de debut :- 5 mars 2007 09:42

Date de fin :- 28 novembre 2011 13:26

< >

+ (direct subAU/total subAU) - xmlID

Ouvrir dossier AU/OG

Recharger groupe d'objets

Messagerie de Christelle Noulet, archiviste aux Archives n

42

26

Sauver unité d'archive

Sauver groupe d'objets

Resip: exportation

The screenshot displays the Resip application window. On the left, a tree view titled 'Arbre du SIP (23746 archiveUnit/2382 dog/35721 bdo/0 pdo)' shows a hierarchy of folders. A blue arrow points from the 'Export ?' menu item in the top navigation bar to the 'Export ?' menu in the foreground. The foreground menu is open, showing several options: 'Exporter le SIP SEDA...', 'Exporter le manifest Xml SEDA...', 'Exporter la hiérarchie sur disque...', 'Exporter la hiérarchie simplifiée et le csv sur disque...', 'Exporter la hiérarchie simplifiée et le csv en zip...', and 'Exporter le csv des métadonnées...'. The 'Exporter le SIP SEDA...' option is highlighted. The background window shows details for an 'Unité d'archive - ID10 ...' with fields for 'Niveau de description', 'Titre', 'Description', 'Date de début', and 'Date de fin'.


When processing is done...
Exportation!


Two AIPs for a single mailbox: View in the digital repository (search results)

 Messagerie électronique de [redacted], directeur adjoint du cabinet de [redacted] (PST)
20190460_2_1 8 septembre 2014 - 10 février 2016

> [membre de cabinet ministériel](#) [culture](#) [Valois \(rue de\)](#)

Produit par : FRAN_NP_052021 - France. Ministère de la Culture et de la communication. Cabinet de [redacted]

[Plus de détails](#) 

 Messagerie électronique de [redacted], directeur adjoint du cabinet de [redacted]
20190460_1_1 8 septembre 2014 - 10 février 2016

> [membre de cabinet ministériel](#) [culture](#) [Valois \(rue de\)](#)

Produit par : FRAN_NP_052021 - France. Ministère de la Culture et de la communication. Cabinet de [redacted]

[Plus de détails](#) 

Raw PST export AIP


Culture

Messagerie électronique de [redacted], directeur adjoint du cabinet de [redacted] (PST)

20190460_2_1 Communicable à partir du 10 février 2066

8 septembre 2014 → 10 février 2016

membre de cabinet ministériel culture Valois (rue de)



Culture

Messagerie électronique de [redacted] directeur adjoint du cabinet de [redacted] (PST)

20190460_2_1 Communicable à partir du 10 février 2066

8 septembre 2014 → 10 février 2016

membre de cabinet ministériel culture Valois (rue de)

Plus d'actions ▾

> Informations *

> Description

▼ Fichiers liés

Usages	Numéro de version	Nom	Format	Taille (octet)	Dernière modification	Téléchargement
Version originale	1	archive.pst	Microsoft Outlook Personal Folders (Unicode)	1485382656	26 avril 2018	

> Règles de gestion

ReSIP-processed AIP

Culture

Messagerie électronique de [redacted], directeur adjoint du cabinet de [redacted]

Messagerie électronique de [redacted], directeur adjoint du cabinet de [redacted]

20190460_1_1 Communicable à partir du 10 février 2066

8 septembre 2014 → 10 février 2016

membre de cabinet ministériel culture Valois (rue de)

Plus d'actions ▾

Messagerie électronique de [redacted]
directeur adjoint du cabinet de [redacted]

ReSIP-processed AIP

Messagerie électronique de [redacted], directeur adjoint du cabinet de [redacted]

Boîte de réception

- Parlement
 - TR: PLF 2015 - Réponses aux QP III-3.4/IX-9.3/V-5.13/X-11.1,2et3
 - TR: Calendrier budgétaire 2015
 - TR: PLF 2015 -Liste des QP sensibles
 - [redacted] - AN - I.15 - BeIN Sport - PLF 2015.odt
 - [redacted] - AN - I.08 - Modernisation de la TNT - PLF 2015.odt
- Photographie
 - CNP
 - Photojournalisme
 - TR: Projet de décret - salaire minimum des photojournalistes pigistes
 - Re: TR: Projet de décret - salaire minimum des photojournalistes pigistes
 - Fwd: Note photojournalisme pour la réunion de 15h
 - [liste.afp] [redacted]
 - annonce la création d'un conseil national de la photo
 - [liste.afp] [redacted]
 - annonce la création d'un conseil national de la photo
 - Réforme sécurité sociale
- Presse et audiovisuel
- Protection sociale

Boîte de réception

- Parlement
 - TR: PLF 2015 - Réponses aux QP III-3.4/IX-9.3/V-5.13/X-11.1,2et3
 - TR: Calendrier budgétaire 2015
 - TR: PLF 2015 -Liste des QP sensibles
 - [redacted] - AN - I.15 - BeIN Sport - PLF 2015.odt
 - [redacted] - AN - I.08 - Modernisation de la TNT - PLF 2015.odt
- Photographie
 - CNP
 - Photojournalisme
 - TR: Projet de décret - salaire minimum des photojournalistes pigistes
 - Re: TR: Projet de décret - salaire minimum des photojournalistes pigistes
 - Fwd: Note photojournalisme pour la réunion de 15h
 - [liste.afp] [redacted]
 - annonce la création d'un conseil national de la photo
 - [liste.afp] [redacted]
 - annonce la création d'un conseil national de la photo
 - Réforme sécurité sociale
- Presse et audiovisuel
- Protection sociale

Messagerie électronique de [redacted], directeur adjoint du cabinet de [redacted]

20190460_1_1 **Communicable à partir du 10 février 2066**

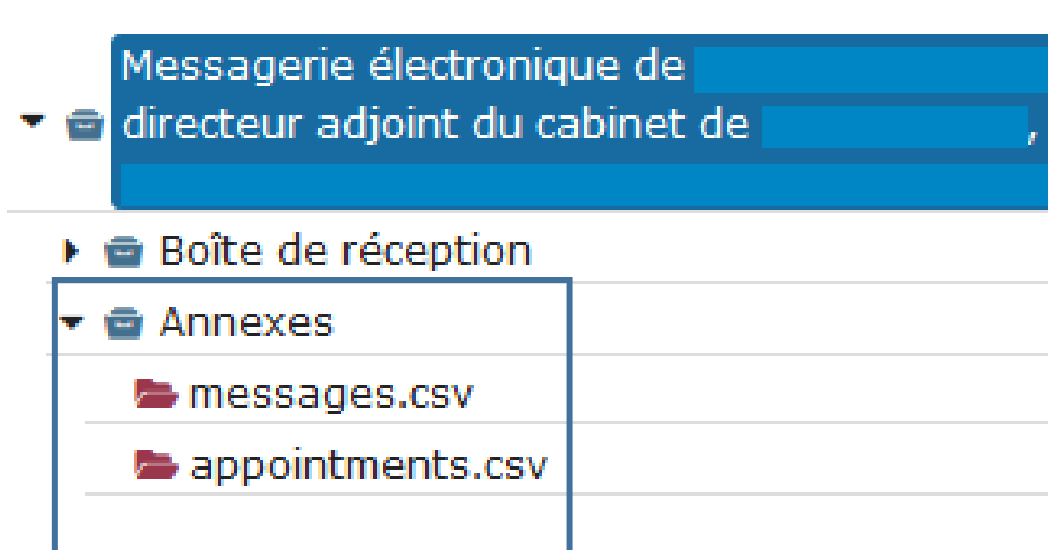
8 septembre 2014 → 10 février 2016

membre de cabinet ministériel culture Valois (rue de)

Plus d'actions

- Informations *
- Description
- Fichiers liés
- Règles de gestion

ReSIP-processed AIP



3. Further considerations and prospects



Limitations

- Much of the process relies on the mission archivists, who handle the exportation process from the client.
- Strategy relies on Resip... Importation errors occur often.
- Require being careful for big mailboxes => several exports => several SIPs.
- Tree structure rendering is not always suitable.

Experiments

- First request for access on a mailbox! But...
 - Restrictions on access => need for a specific authorization => investigation to accept or refuse (2 months delay)
 - Haven't been ingested yet
 - Import didn't work in ReSIP
 - Impossibility to plan processing and ingest
- 1 mailbox but 4 PST (9 Gb) but how many messages?
 - Ratom => SQLite file
 - 95 000 messages, 45 000 attachments
 - Half the mailbox is likely to answer
 - Impossibility to check precisely the relevance of the results and the presence of sensitive information
 - Epadd+ (with Emailchemy trial license)
 - More or less same results (but some messages were not imported)
 - Access to the attachments
 - But same conclusion as Ratom

Reflections

- For this request, we couldn't go further in practice but it raised some questions for possible « regular cases ».
- What is the most effective way to perform sensitivity checking and relevance checking?
 - On the whole mailbox? How must it be rendered?
 - On a finding aid as a csv file?
 - Directly through the digital repository?
- How to actually give researchers access to the mailbox? The entire mailbox? A collection of messages? In what form and what format?
- When the archive service refuses access, the researcher can appeal against the decision to the CADA (Commission for administrative records access). Difficulty to process mailboxes has already been taken into account to reject the request.

Conclusion

- Lot of questions, but no answers yet regarding access...
- Our strategy = precaution
- Leaves space to consider potential future answers while preserving mailboxes and their messages
- Can evolve...!
- But we need a lot of time to experiment with tools

Thank you for your attention!
Time for Q & A session

