



# Générateur de SIP pour un arbre de positionnement ou un plan de classement

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.

Version du 15/06/2018



# SOMMAIRE



**01** Présentation du générateur

**02** Établissement du fichier d'import

**03** Fonctionnement simple du générateur

**04** Annexe : Rappels sur le SEDA 2.0.



# Documents de référence



- Vitam. Modèle de données
- Vitam. Modèle de workflow
- Vitam. Arbre de positionnement et plan de classement
- Vitam. Manuel utilisateur



# 1/ Présentation du générateur

--

# Avertissements

- Les spécifications et le développement de cette extension du générateur ont été très rapides (1/2 journée de travail pour un développeur) => il n'est pas possible de considérer ce produit autrement que comme un démonstrateur ou un prototype
- Le fichier source qui permet de créer l'arbre de positionnement/le plan de classement a été construit par le MEAE pour la reprise de données des plans de classements du ministère dans Diplomatie, sans que l'équipe Vitam procède à des modifications sur ce fichier => il n'a pas été créé pour être alimenté par des archivistes, mais par des informaticiens

# Objectifs de l'outil

- Faciliter la génération des jeux de tests pour les membres de l'équipe Vitam, les ministères porteurs et les partenaires
- Générer automatiquement, à partir d'un tableur, des SIP correspondant à des arbres de positionnement/cadres de classement ou à des plans de classement
  - Compatibles NF Z44-022 et standard SEDA v2.1 (conforme schéma .xsd du SEDA 2.1.)
  - Conformes au document de spécification des SIP propre à la solution logicielle Vitam
  - Sans nécessairement avoir besoin d'utiliser un éditeur xml pour créer le bordereau

# Fonctionnement global de l'outil



- Le générateur de SIP pour des arbres de positionnement et des plans de classement fonctionne comme le générateur de SIP « standard »
- La seule différence réside dans la source des données :
  - générateur de SIP « standard » : la source des données est un répertoire Windows contenant une arborescence de répertoires et des fichiers
  - générateur de SIP pour des arbres de positionnement et des plans de classement : la source des données est un fichier .csv contenant l'ensemble des informations à traiter



## 2/ Établissement du fichier d'import

--





## 2.1./ Constitution du fichier source



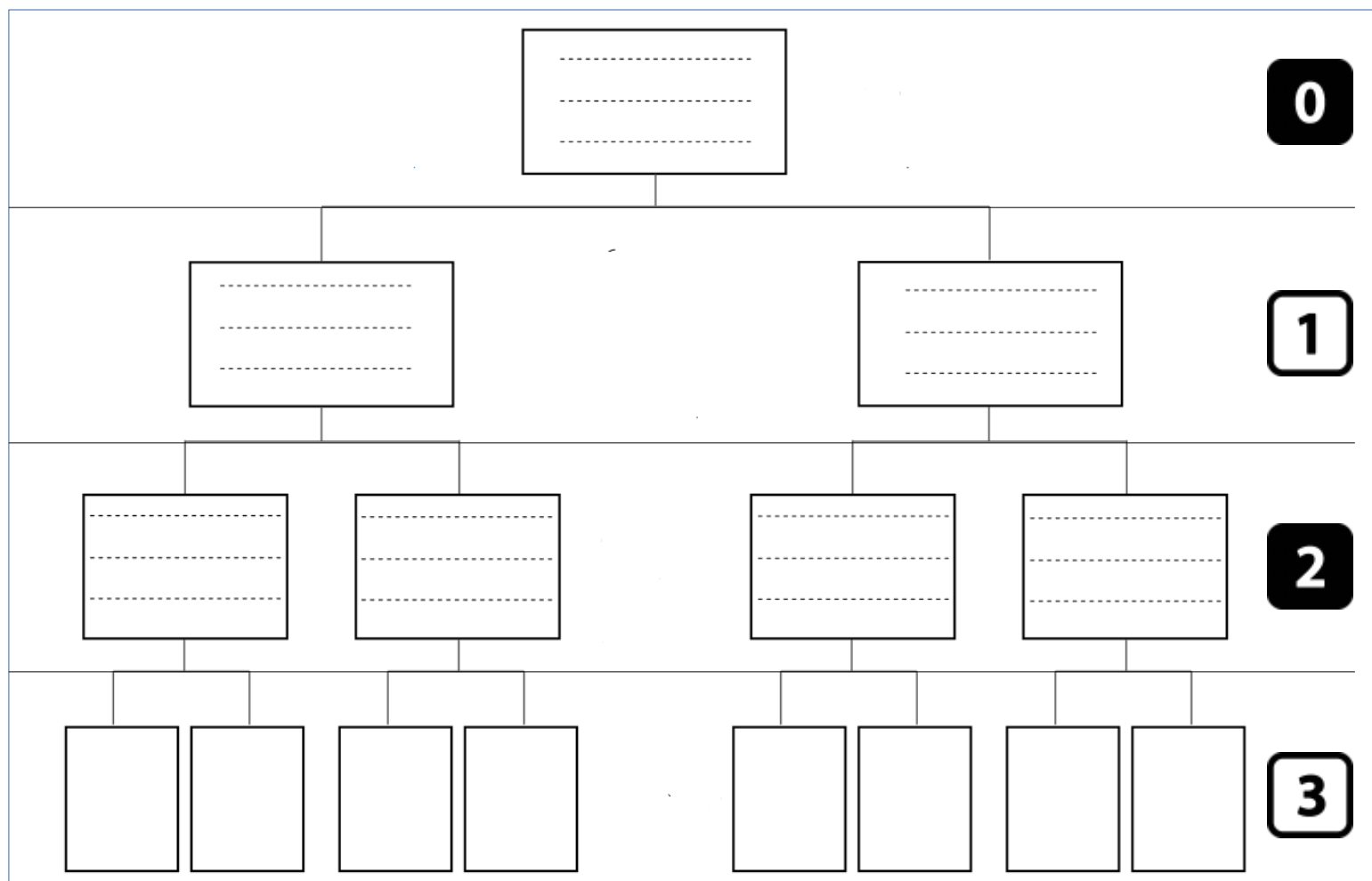
# Présentation du fichier source

- Le fichier source prend la forme d'un tableur composé de 7 colonnes :
  - id : numéro séquentiel
  - nom : intitulé de l'unité de description
  - observations : champ libre
  - cote : identifiant métier propre à l'unité de description (suffixe)
  - série : identifiant métier de l'unité de description « mère » (préfixe)
  - rang : à chaque niveau, ordre des unités de description
  - niveau : profondeur par rapport à la racine de l'arbre/plan
- Nota bene* : l'ordre des colonnes dans le fichier lui-même ne doit pas être modifié

	A	B	C	D	E	F	G
1	Id	nom	<u>observ</u>	cote	<u>serie</u>	rang	niveau
2	1	Cabinet de Michel Mercier			CAB.	1	0
3	2	Fonctionnement du cabinet		1.	CAB.	1	1
4	3	Travail gouvernemental		2.	CAB.	2	1
5	4	Communication		3.	CAB.	3	1
6	5	Discours du ministre		1.	CAB.3.	1	2
7	6	Affaires civiles et sceau		4.	CAB.	4	1
8	7	Affaires criminelles et grâce		5.	CAB.	5	1
9	8	Protection judiciaire de la jeunesse		6.	CAB.	6	1
10	9	Administration pénitentiaire		7.	CAB.	7	1
11	10	Professions judiciaires		8.	CAB.	8	1
12	11	Fonctionnement du ministère		9.	CAB.	9	1
13							
14							

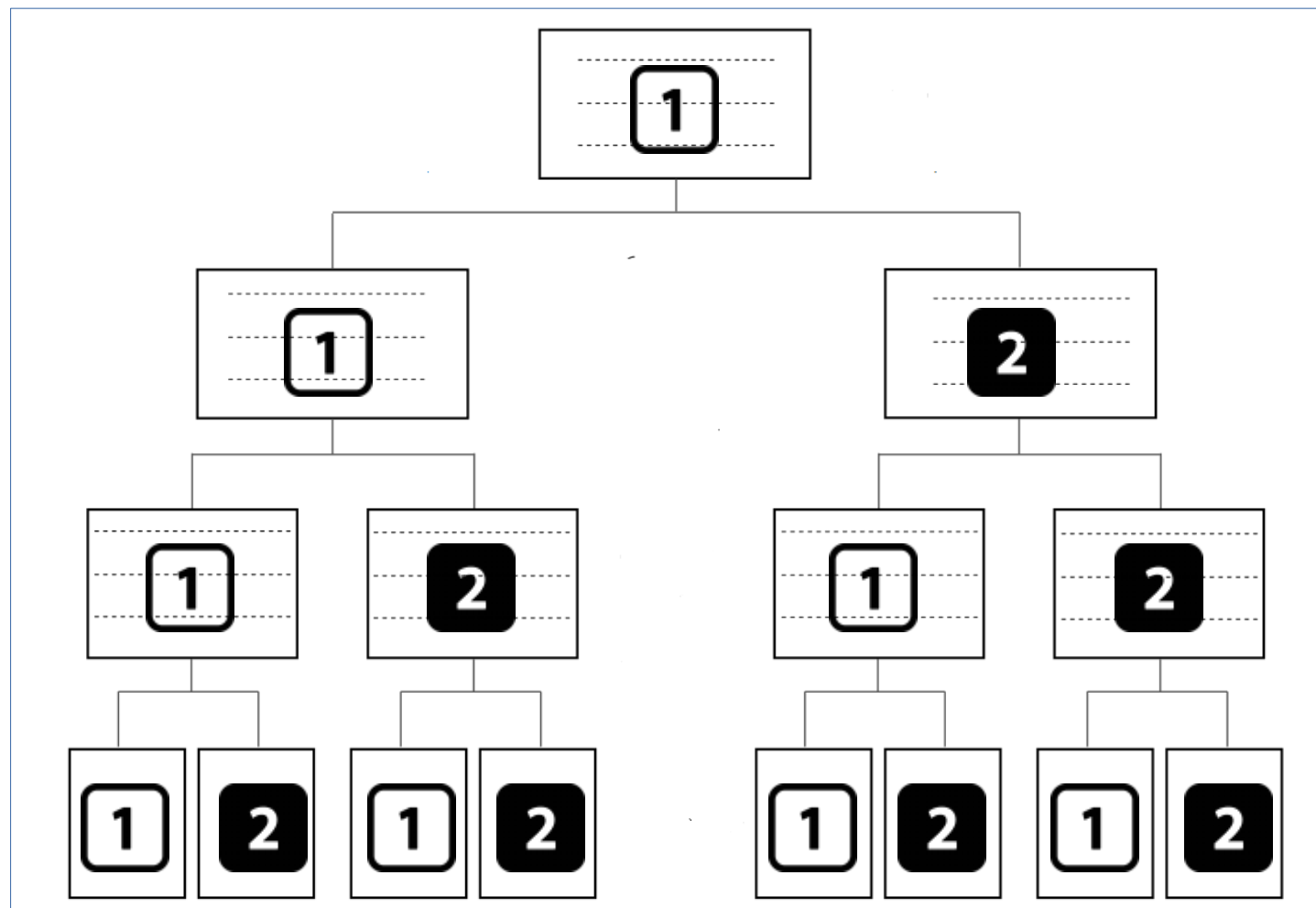
# Comment remplir la colonne « Niveau » ?

La colonne  
« niveau » permet  
d'indiquer la  
« profondeur » de  
l'unité de  
description par  
rapport à la racine  
de l'arbre/du plan  
Elle doit  
obligatoirement  
être renseignée  
La racine a pour  
niveau « 0 »



# Comment remplir la colonne « Rang » ?

La colonne « rang » permet d'indiquer l'ordre des unités de descriptions au sein d'un même niveau et d'une même branche de l'arbre/du plan. Elle doit obligatoirement être renseignée.



# Comment remplir les colonnes « série » et « cote » ?

- Les colonnes « série » et « cote » sont complémentaires et permettent de définir l'identifiant de l'unité de description (champ OriginatingAgencyArchiveUnitIdentifier dans le SEDA)
- La colonne « série » doit obligatoirement être renseignée
- Prenons quelques exemples :
  - aux Archives nationales :
    - Série F/ : dans le fichier, F/ sera la série et il n'y aura pas de cote
    - Sous-série F/1 : dans le fichier, F/ sera la série et 1 la cote
    - Sous-série F/1c : dans le fichier, F/1 sera la série et c la cote
    - Sous-série F/1cVII : dans le fichier, F/1c sera la série et VII la cote
  - Aux Archives départementales, pour le cadre de classement des archives communales
    - Série F/ : dans le fichier, F/ sera la série et il n'y aura pas de cote
    - Sous-série 1F : dans le fichier F/ sera la série et 1F la cote

# Points d'attention



- Ne pas utiliser de caractère utilisé comme un séparateur pour la création d'un fichier .csv, notamment les « ; »
- Vérifier qu'il n'y ait pas d'espaces vides dans les colonnes id, rang et niveau, ainsi que dans les autres colonnes, s'ils ne sont pas souhaités
- Éviter les descriptions trop longues dans le champ « Description », sa taille étant limitée

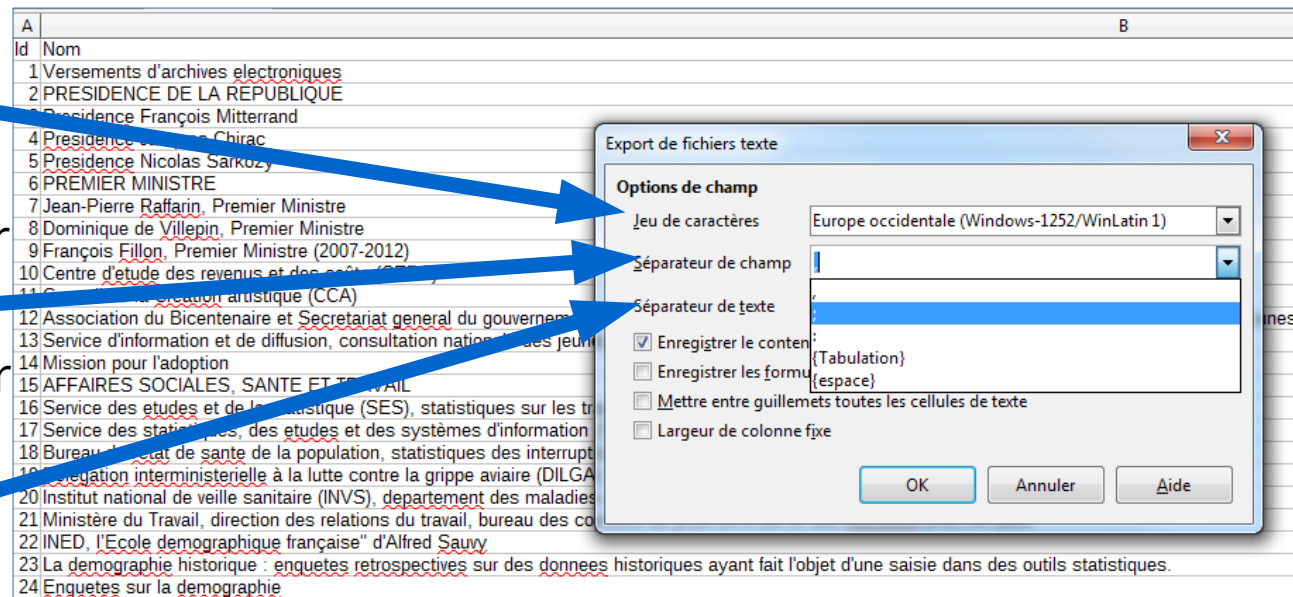


## 2.2./ Conversion au format .csv

--

# Étape 1 – Convertir le fichier au format .csv

- Ouvrir le tableur
- Sélectionner « Enregistrer sous » et choisir le format « CSV »
- Dans la fenêtre de dialogue :
  - Choisir comme jeu de caractères « Europe occidentale (Windows-1252/WinLatin1) »
  - choisir comme séparateur de champ « ; »
  - vider le champ séparateur de texte
- Procéder à l'enregistrement





## Étape 2 – Supprimer la dernière ligne et régler les problèmes d'encodage

- Ouvrir le fichier .csv dans un éditeur de texte (Bloc Notes, WordPad ou Notepad++)
- Supprimer la dernière ligne qui est vide
- Modifier l'encodage :
  - Dans Notepad ++, sélectionner « Encodage » puis « Encoder en UTF-8 »
  - Dans les autres outils, sélectionner « Enregistrer sous » et, dans la fenêtre de dialogue, la valeur « UTF-8 » dans le champ « Encodage »
- Enregistrer le fichier
- *Nota bene* : la version actuelle du produit n'est pas totalement satisfaisante en termes d'encodage des caractères. Il peut être nécessaire de modifier à la main les caractères accentués dans le manifeste .xml final

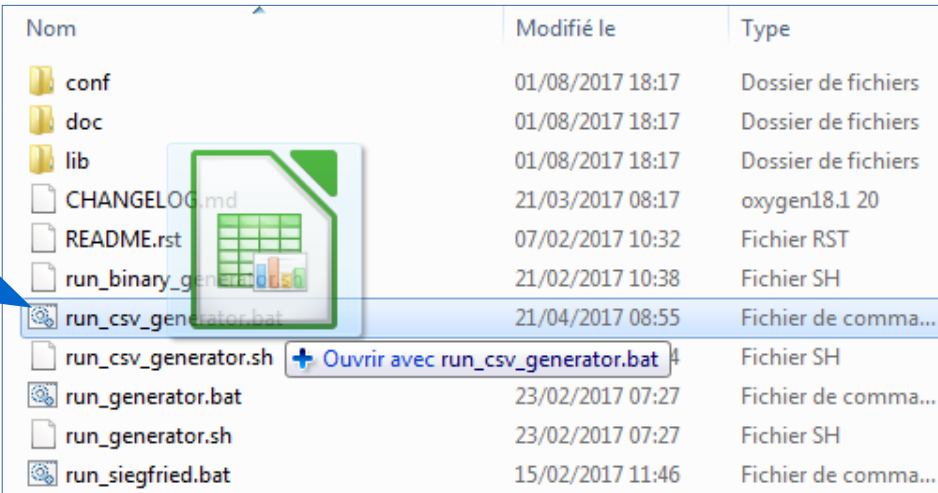


### 3/ Fonctionnement du générateur

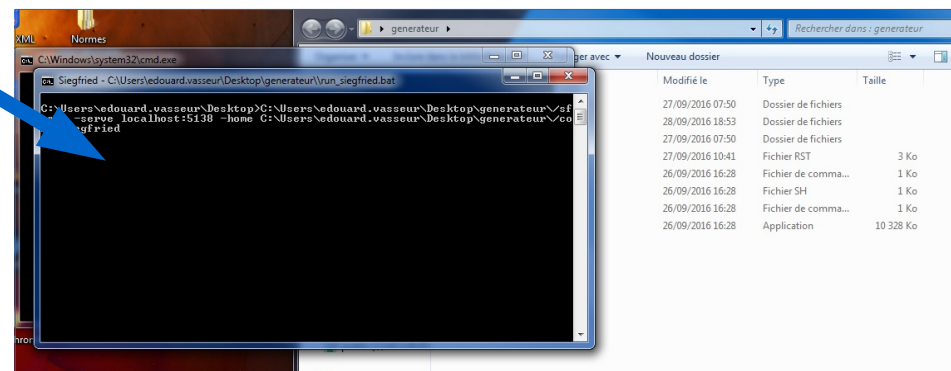
--

# Étape 1 - Lancement du générateur

- Ouvrir le répertoire contenant le générateur
- Faire glisser le fichier.csv vers le fichier « run\_csv\_generator.bat »
  - Deux consoles s’ouvrent. La 2<sup>e</sup> console trace les erreurs survenues pendant l’opération. Elle se ferme à la fin de l’opération en tapant sur la touche « entrée »

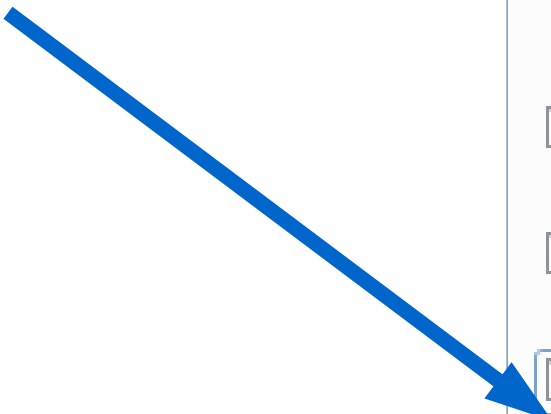


Nom	Modifié le	Type
conf	01/08/2017 18:17	Dossier de fichiers
doc	01/08/2017 18:17	Dossier de fichiers
lib	01/08/2017 18:17	Dossier de fichiers
CHANGELOG.md	21/03/2017 08:17	oxygen18.1 20
README.rst	07/02/2017 10:32	Fichier RST
run_binary_generator.bat	21/02/2017 10:38	Fichier SH
run_csv_generator.bat	21/04/2017 08:55	Fichier de comma...
run_csv_generator.sh		+ Ouvrir avec run_csv_generator.bat Fichier SH
run_generator.bat	23/02/2017 07:27	Fichier de comma...
run_generator.sh	23/02/2017 07:27	Fichier SH
run_siegfried.bat	15/02/2017 11:46	Fichier de comma...



# Étape 2 – où retrouver le SIP généré ?

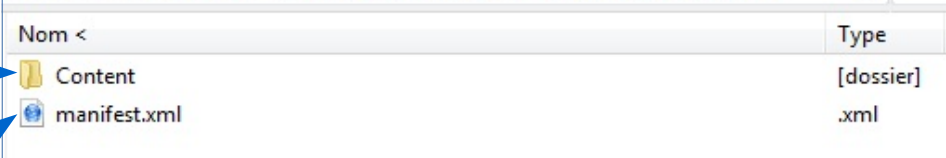
Le fichier zip SIP a été créé dans le répertoire contenant le générateur



conf	01/08/2017 18:17
doc	01/08/2017 18:17
lib	01/08/2017 18:17
CHANGELOG.md	21/03/2017 08:17
README.rst	07/02/2017 10:32
run_binary_generator.sh	21/02/2017 10:38
run_csv_generator.bat	21/04/2017 08:55
run_csv_generator.sh	06/04/2017 14:34
run_generator.bat	23/02/2017 07:27
run_generator.sh	23/02/2017 07:27
run_siegfried.bat	15/02/2017 11:46
SIP-20170919091330.zip	19/09/2017 09:13

# Étape 3 – Consulter le SIP

- Ouvrir le zip et constater qu'il contient bien :
  - Un répertoire Content vide en raison de l'absence de fichiers à archiver dans le SIP
  - Un fichier manifest.xml qui reprend le contenu du tableur (et donc de l'arbre ou du plan de classement) sous forme de bordereau SEDA



Nom <	Type
Content	[dossier]
manifest.xml	.xml

NB : le bordereau peut être extrait pour enrichissement (ex. avec un éditeur xml) et réinjecté dans le SIP ensuite (en supprimant la version d'origine)

Le SIP est prêt !