



Guide d'application

Table des matières

1	Préambule	2
2	Configurations système requises.....	3
3	Problèmes en suspens / Commentaires / Souhaits de modifications	3
4	Installation.....	4
4.1	Installation de SIP-Val.....	4
4.2	Installation des applications tierces	4
5	Configuration de SIP-Val	7
5.1	Propriétés du fichier de configuration « sipvalidator.conf.xml »	7
5.2	« DROID_SignatureFile_V59-kc.xml » et « SignatureFile.xsd »	10
5.3	« report.xsl »	10
6	cmd_Script-Example.....	11
6.1	« _SIP-Val_demarrage.cmd »	11
6.2	« _SIP-Val_demarrage_+3c_+3d.cmd »	11
6.3	« commande_manuelle_SIP-Val.cmd »	11
7	Modifications pour la version française	12
8	Exécuter la validation.....	13
8.1	Exécuter en utilisant le script de validation	13
8.2	Démarrer la validation manuellement	14
9	Évaluer la validation.....	16
10	Description des modules de validation et des étapes de validation.....	18
10.1	Module 1: Consistance paquet et XML	18
10.2	Module 2: Consistance du fichier	20
10.3	Module 3: Consistance du format de fichier et de la date	20
11	Droits d'auteur.....	21

1 Préambule

SIP-Val est une application java de console pour la validation de la configuration et du contenu de ce qu'on appelle un paquet d'informations à verser (SIP ; *Submission Information Package*) pour le versement d'informations numériques. Cette application est sous licence GPL3+ et elle est mise à disposition du public par le CECO en Open Source. SIP-Val remplit les exigences décrites ci-après.

Exigences fonctionnelles:

SIP-Val lit un SIP et valide les points suivants de la spécification SIP des AFS de septembre 2009 dans la version 1.0¹ et de la spécification SIP, eCH 0160, de novembre 2012, dans la version 1.0². Cela correspond à la spécification SIP des AFS de mars 2012 dans la version 4.0³:

Étape de validation	Nom de l'étape (Stepname)
1a (abandon en cas d'erreur)	lisibilité
1b (abandon en cas d'erreur)	validation de la structure des données
1c (abandon en cas d'erreur)	nommage de la structure et des fichiers
1d (abandon en cas d'erreur)	validation du schéma metadata.xml
1e	déterminer le type du SIP"
1f	existence des données primaires
2a	données primaires manquant
2b	validation de la somme de contrôle
2c	données primaires supplémentaire
2d	description des données primaires GEVER
3a	détection des formats
3b	formats supplémentaire
3c (facultatif)	validation de formats
3d (facultatif)	validation de période

Les résultats (y compris les messages d'incohérences ou d'erreurs) de chaque étape sont écrites dans un fichier de log.

Le résultat de la validation globale (SIP correct/erroné) est aussi délivré et visible dans le statut *exit* du programme, afin que la validation puisse être liée dans la chaîne de traitement automatisée. Le statut *exit* peut prendre les valeurs suivantes:

- 0 tout est OK, valide
- 1 informations erronées
- 2 échec de la validation, invalide

Les étapes de validation / les vérifications sont exécutées les unes après les autres. Lorsque cela est possible, la validation continue en cas d'erreur afin de réduire le nombre de cycles de correction. Seules les vérifications 1a) à 1d) dans le module 1 «Consistance paquet et XML» sont des erreurs qui entraînent directement l'interruption du programme.

¹ La spécification peut être téléchargée sur la page «L'interface de versement numérique» du site des AFS: <http://www.bar.admin.ch/themen/00876/00877/index.html?lang=fr>.

² La spécification peut être téléchargée à partir du site Web de eCH: <http://www.ech.ch/vechweb/page?p=dossier&documentNumber=eCH-0160>.

³ La spécification peut être téléchargée sur la page «L'interface de versement numérique» du site des AFS: <http://www.bar.admin.ch/themen/00876/00877/01561/index.html?lang=fr>.

Exigences non fonctionnelles:

Pour des tâches particulières (extraction ZIP, validation XML, reconnaissance du format), des programmes externes ou des frameworks Java appropriés sont utilisés.

L'application est constituée de manière modulaire, de sorte que la mise en place d'un ou plusieurs module(s) de validation supplémentaire(s) nécessite moins d'effort.

Le schéma déposé dans le SIP est utilisé pour la validation des schémas (validation dynamique des schémas).

La sortie du programme et le fichier log permettent d'utiliser l'outil dans la chaîne de processus, c'est-à-dire que le résultat de chaque validation peut être facilement déterminé.

La sortie de la console se limite à la description des étapes de validation (y compris le module de test correspondant), les résultats valides ou invalides, le résultat global, ainsi que les chemins vers le SIP et le fichier log généré. Toutes les informations additionnelles sont spécifiées dans le fichier log.

2 Configurations système requises

Microsoft Windows 98, ME, NT, 2000, XP, VISTA ou Windows 7

Au moins 128 MB de RAM (512 MB ou plus recommandé)

Au moins 20 GB d'espace sur le disque dur

Java Runtime Environment (JRE), version 6 ou plus récente⁴ (téléchargement: <http://www.java.com>)

3 Problèmes en suspens / Commentaires / Souhaits de modifications

Pour les points ouverts (erreurs, bogues, souhaits de modifications) prière de consulter la plate-forme de développement GitHub: <https://github.com/KOST-CECO/SIP-Val/issues>

Cette liste peut être complétée par tous et est traitée par l'équipe de développement.

⁴ On peut vérifier la version actuelle installée avec la commande `java -version` dans la fenêtre cmd.

4 Installation

4.1 Installation de SIP-Val

1	Créer le répertoire SIP-Val.
2	Télécharger SIP-Val et l'extraire dans le répertoire SIP-Val.

The first screenshot shows a Windows Explorer window with the address bar set to 'Computer > System (C:) > Software > KOST-Val'. The file list shows a new folder named 'SIP-Val' of type 'Dateiordner'. The second screenshot shows the same Explorer window with the address bar set to '« System (C:) > Software > KOST-Val > SIP-Val'. A green box highlights the contents of the 'SIP-Val.zip' file being extracted into the 'SIP-Val' folder. The extracted files are: 'cmd_Script-Example' (Dateiordner), 'configuration' (Dateiordner), 'doc' (Dateiordner), 'license' (Dateiordner), 'fr_sipvalidator.jar' (Executable Jar File), and 'sipvalidator.jar' (Executable Jar File).

4.2 Installation des applications tierces

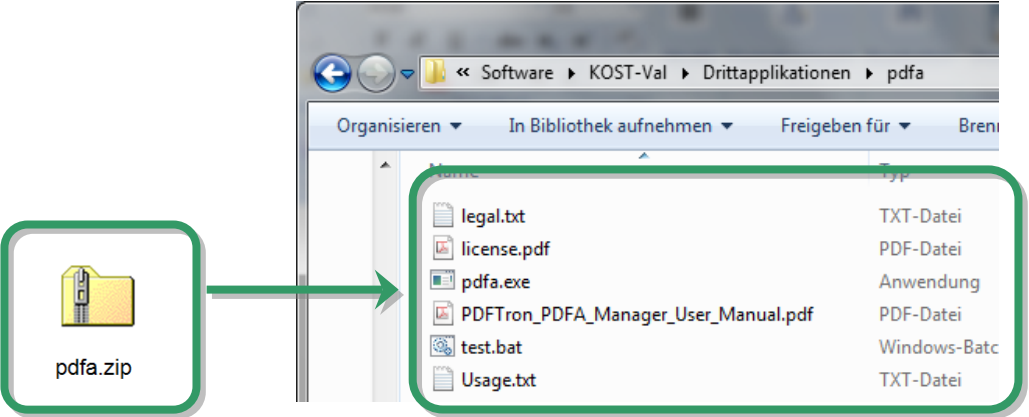
4.2.1 Installation de SIARD-Val

1	Télécharger la variante de la ligne de commande de SIARD-Val (ZIP-File) sur http://kost-ceco.ch/cms/index.php?siard_val_fr .
2	Extraire SIARD-Val dans le répertoire souhaité ou bien dans le répertoire prévu SIARD-Val.

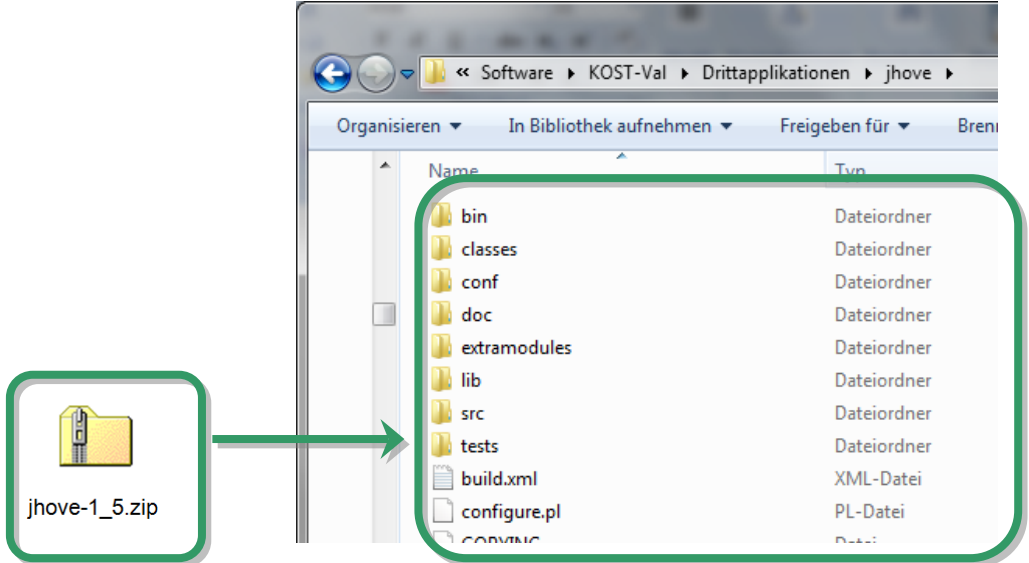
The screenshot shows a Windows Explorer window with the address bar set to '« System (C:) > Software > KOST-Val > SIARD-Val'. A green box highlights the contents of the 'SIARD-Val.zip' file being extracted into the 'SIARD-Val' folder. The extracted files are: 'cmd_Script-Example' (Dateiordner), 'configuration' (Dateiordner), 'doc' (Dateiordner), 'license' (Dateiordner), 'fr_siard-val.jar' (Executable Jar File), and 'siard-val.jar' (Executable Jar File).

Note:
Si un autre répertoire est choisi, le fichier de configuration "sipvalidator.conf.xml" doit être adapté. Voir le chapitre 5.

4.2.2 Installation de PDFTRON (PDF/A Manager Trial Version)

1	Télécharger la variante de la ligne de commande du PDF/A Manager (fichier ZIP) sur http://www.pdftron.com/downloads.html .
2	<p>Extraire PDF/A Manager dans le répertoire souhaité ou bien dans le répertoire prévu Drittapplikationen\pdfa.</p>  <p>Note: Si un autre répertoire est choisi, le fichier de configuration "sipvalidator.conf.xml" doit être adapté. Voir le chapitre 5.</p>

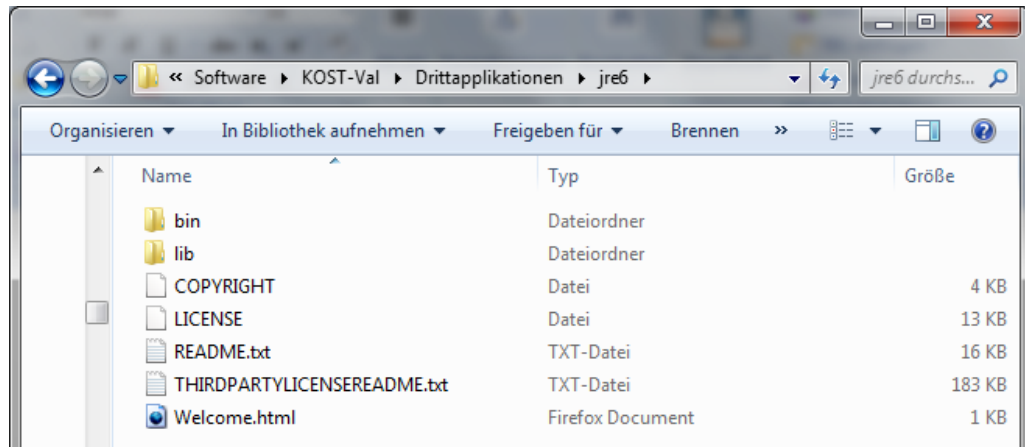
4.2.3 Installation de JHOVE (version 1.5)

1	Télécharger l'application JHOVE (fichier ZIP) sur http://sourceforge.net/projects/jhove/ .
2	<p>Extraire JHOVE dans le répertoire souhaité ou bien dans le répertoire prévu Drittapplikationen\jhove.</p>  <p>Note: Si un autre répertoire est choisi, le fichier de configuration "sipvalidator.conf.xml" doit être adapté. Voir le chapitre 5.</p>

4.2.4 Installation de Java (version 6 ou plus récente)

1

Si aucune version 6 ou plus récente de Java Runtime Environment (JRE)⁵ n'est déjà installée, jre6 doit être téléchargé sur <http://www.java.com> et être enregistré dans Driitapplikationen\jre6.



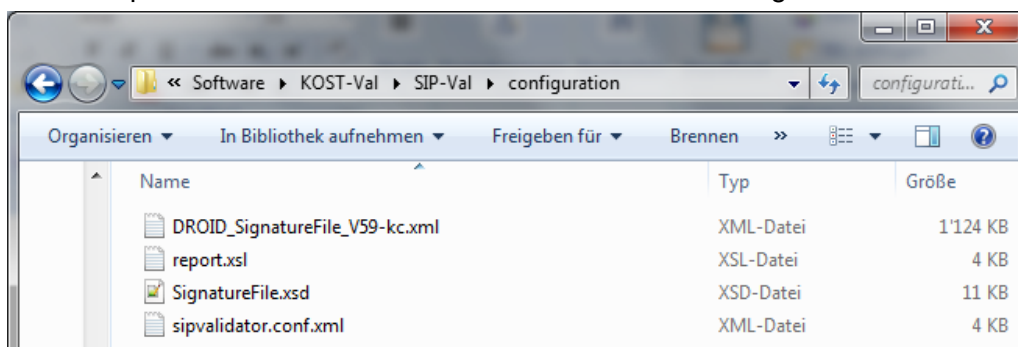
Note:

Si jre version 6 ou plus récente est déjà installée, on peut renoncer de l'enregistrer à nouveau. A noter que vous devez modifier le script cmd comme décrit au chapitre 6.

⁵ Vérifier la version avec la commande `java -version` dans la fenêtre cmd.

5 Configuration de SIP-Val

Toutes les configurations de SIP-Val peuvent être effectuées dans le fichier de configuration « sipvalidator.conf.xml » dans le sous-dossier « configuration ».



Les fichiers « DROID_SignatureFile_V59-kc.xml », « SignatureFile.xsd » et « report.xml » sont aussi enregistrés dans le dossier « configuration ». Ceux-ci sont décrits dans le chapitre 5.2.

5.1 Propriétés du fichier de configuration « sipvalidator.conf.xml »

balise xml	Description
<allowedlengthofpaths>	Nombre maximal de caractères autorisé dans la longueur du chemin
<allowedlengthoffiles>	Nombre maximal de caractères autorisé dans le nom du fichier
<allowsipname>	Spécifications pour la composition du nom SIP
<pathtopdftronexe>	Chemin d'accès à PDFTRON (voir aussi chapitre 4.2, Installation des applications tierces)
<pathtosiardvaljar>	Chemin d'accès à SIARD-Val (voir aussi chapitre 4.2, Installation des applications tierces)
<pathtojhovejar>	Chemin d'accès à JHOVE (voir aussi chapitre 4.2, Installation des applications tierces)
<pathtojhoveconfig>	Chemin d'accès au fichier de configuration de JHOVE (voir aussi chapitre 4.2, Installation des applications tierces)
<pathtoworkdir>	Chemin d'accès au répertoire de travail temporaire du validateur SIP
<pathtodroidssignature>	Chemin d'accès au fichier de signature DROID utilisé
<allowedversionbar1>	Indiquant si la version 1 du AFS-SIP est autorisé
<allowedversionbar4ech1>	Indiquant si la version 4 du AFS-SIP est autorisé respectivement eCH-0160 version 1
<allowedformats>	Récapitulatif des formats de fichiers autorisés, se composant de sous-éléments avec des informations détaillées
<validatedformats>	Récapitulatif des formats de fichiers à valider, se composant de sous-éléments avec des informations détaillées Pour l'étape de validation optionnelle 3c, l'application tierce à utiliser peut être attribuée dans la sous-balise <validateur> aux différents formats (JHOVE, PDFTRON ou SIARDVAL).

5.1.1 <allowedlengthofpaths>

Nombre maximal de caractères autorisé dans la longueur du chemin La valeur initiale est 179 et correspond à la restriction du SIP des AFS: moins de 180 caractères.

5.1.2 <allowedlengthoffiles>

Nombre maximal de caractères autorisé dans le nom du fichier. La valeur initiale est **40** et correspond à la restriction du SIP des AFS (version 1): moins de 40 caractères. En validant la version 4 du SIP, il n'existe aucune restriction, donc ce champ ne concerne que la version 1.

5.1.3 <allowsipname>

Spécifications pour la composition du nom du SIP. La valeur initiale (expression Regex) est **SIP_[1-2][0-9]{3}[0-1][0-9][0-3][0-9]_\\w{3}** et correspond à la restriction du SIP des AFS: SIP_[dateversement]_[Nom du service versant](_[Référence]), pour lequel la date doit être saisie au format AAAAMMJJ et la référence peut être utilisée de manière optionnelle.

5.1.4 <pathtopdftronexe>

Chemin d'accès à PDFTRON. La valeur initiale est **Drittapplikationen \pdfa\pdfa.exe**. Le chemin absolu peut être saisi ou le chemin relatif comme dans la valeur initiale (du point de vue de démarrage.cmd ou démarrage_+3c_+3d.cmd).

5.1.5 <pathtosiardvaljar>

Chemin d'accès à SIARD-Val. La valeur initiale est **SIARD-Val\siard-val.jar**. Le chemin absolu peut être saisi ou le chemin relatif comme dans la valeur initiale (du point de vue de démarrage.cmd ou démarrage_+3c_+3d.cmd).

5.1.6 <pathtojhovejar>

Chemin d'accès à JHOVE. La valeur initiale est **Drittapplikationen \jhove\bin\JhoveApp.jar**. Le chemin absolu peut être saisi ou le chemin relatif comme dans la valeur initiale (du point de vue de démarrage.cmd ou démarrage_+3c_+3d.cmd).

5.1.7 <pathtojhoveconfig>

Chemin d'accès au fichier de configuration de JHOVE. La valeur initiale est **Drittapplikationen\jhove\conf\jhove.conf**. Le chemin absolu peut être saisi ou le chemin relatif comme dans la valeur initiale (du point de vue de démarrage.cmd ou démarrage_+3c_+3d.cmd).

5.1.8 <pathtoworkdir>

Chemin d'accès au répertoire de travail temporaire du SIP-Val. La valeur initiale est **work SIP-Val**. Le chemin absolu peut être saisi ou le chemin relationnel comme dans la valeur initiale (du point de vue de démarrage.cmd ou démarrage_+3c_+3d.cmd). Ce dossier ne doit pas être créé.

5.1.9 <pathtodroidssignature>

Chemin d'accès au fichier de signature DROID utilisé dans le répertoire. La valeur initiale est **SIP-Val\configuration\DROID_SignatureFile_V59-kc.xml**. Des informations supplémentaires sont disponibles au chapitre 5.2.

5.1.10 <allowedversion1>

Indiquant si la version 1 du AFS-SIP est permis (1) ou non (0). La valeur initiale est **1**.

5.1.11 <allowedversion4>

Indiquant si la version 4 du AFS-SIP respectivement eCH-0160 version 1 est permis (1) ou non (0). La valeur initiale est 1.

5.1.12 <allowedformats>

Récapitulatif des formats de fichiers autorisés, se composant de sous-balises avec des informations détaillées. Les valeurs initiales sont visibles dans le tableau suivant. La liste détaillée est conservée intentionnellement, de sorte que les formats non autorisés puissent être facilement supprimés et que le moins possible n'aient à être ajoutés. Les indications concernant l'extension et le PUID sont visibles sur le site de PRONOM (<http://nationalarchives.gov.uk/PRONOM/Default.aspx>).

<puid> </puid>	<extension> </extension>	<comment> </comment>
x-fmt/111	txt	Plain Text file
fmt/353	tif, tiff	Pas de distinction entre les versions
x-fmt/392	jp2, jpeg2, j2c	JPEG2000
fmt/95	pdf	Reconnaissance des principales caractéristiques de PDF/A-1a
fmt/354	pdf	Reconnaissance des principales caractéristiques de PDF/A-1b
fmt/141	wav, wave	Waveform Audio (PCM WAVEFORMAT)
fmt/6	wav, wave	Waveform Audio
x-fmt/18	csv	Comma Separated Values
fmt/161	siard	Reconnaissance des principales caractéristiques de SIARD v1

5.1.13 <validatedformats>

Récapitulatif des formats de fichiers à valider, se composant de sous-balises avec des informations détaillées. Les valeurs initiales sont visibles dans le tableau suivant. La liste détaillée est conservée intentionnellement, de sorte que les formats non autorisés puissent être facilement supprimés et que le moins possible n'aient à être ajoutés. Les indications concernant l'extension et le PUID sont visibles sur le site de PRONOM (<http://nationalarchives.gov.uk/PRONOM/Default.aspx>). Pour l'étape de validation optionnelle 3c, l'application tierce à utiliser peut être attribuée dans la sous-balise <validateur> aux différents formats (JHOVE ou PDFTRON).

<pronomuniqueid> </pronomuniqueid>	<validator> </validator>	<description> </description>
x-fmt/111	-	Plain Text file
fmt/95	PDFTRON	PDF/A-1a
fmt/354	PDFTRON	PDF/A-1b
x-fmt/18	-	Comma Separated Values
fmt/161	SIARDVAL	SIARD
fmt/353	JHOVE	TIFF
x-fmt/392	JHOVE	JPEG2000
fmt/141	JHOVE	Waveform Audio (PCM WAVEFORMAT)
fmt/6	JHOVE	Waveform Audio

5.2 « DROID_SignatureFile_V59-kc.xml » et « SignatureFile.xsd »

Les fichiers « DROID_SignatureFile_V59-kc.xml » et « SignatureFile.xsd » sont enregistrés dans le dossier « configuration ». « SignatureFile.xsd » est le schéma XML pour « DROID_SignatureFile_V59-kc.xml » et définit sa structure.

« DROID_SignatureFile_V59-kc.xml » est nécessaire pour la reconnaissance du format. Il comprend l'état du DROID au 21 mars 2012 et a été en plus adapté par le CECO dans le domaine de la reconnaissance PDF/A⁶.

Afin de pouvoir mettre à jour régulièrement ce SignatureFile, DROID doit être installé séparément et la mise à jour ne pas être empêchée par le fournisseur de technologie d'information. Environ deux fois par an, le CECO rend disponible au téléchargement le SignatureFile DROID actuel avec son adaptation dans le domaine de la reconnaissance PDF/A, en tant qu'alternative.

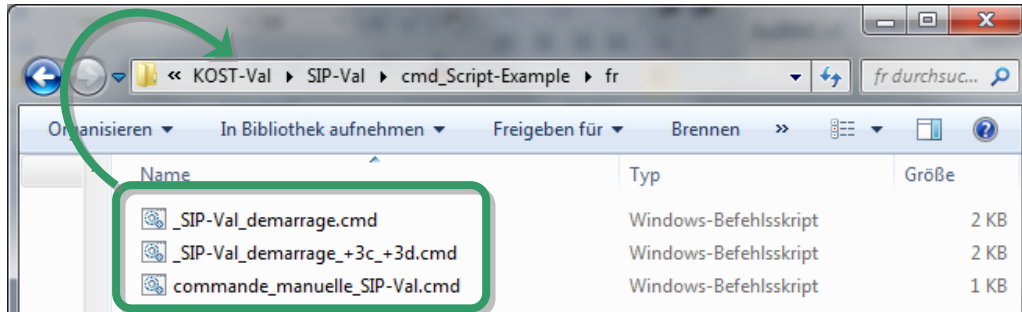
5.3 « report.xsl »

Le fichier « report.xsl » est enregistré dans le dossier « configuration » en tant que modèle de copie. Ce fichier est nécessaire afin que le rapport log PDFTron puisse être représenté de manière lisible dans Internet Explorer.

⁶ De notre point de vue, dans la base de données PRONOM, l'ordre de priorité des formats lors de la reconnaissance PDF n'est pas correctement inscrit. En conséquence la signature interne doit aussi être adaptée. Voir aussi [KorrekturenV52.pdf](#).

6 cmd_Script-Example

Des exemples de scripts de commande sont enregistrés dans le sous-dossier « cmd_Script-Example ». Pour les utiliser il convient de les copier dans le dossier où se trouve le dossier SIP-Val; dans notre exemple, « KOST-Val ».



Les scripts sont brièvement décrits ci-dessous.

6.1 « _SIP-Val_demarrage.cmd »

Ce script valide tous les SIP qui se trouvent dans « KOST-Val\SIPS », indépendamment du fait que ce soient des dossiers ou des fichiers ZIP. Les fichiers LOG sont écrits dans le dossier « KOST-Val\logs ». Ce script utilise également java.exe dans le dossier « Drittapplikationen ». Si jre6 est déjà installé sur l'ordinateur, remplacer « Drittapplikationen\jre6\bin\java.exe » dans le script avec « java ».

6.2 « _SIP-Val_demarrage_+3c_+3d.cmd »

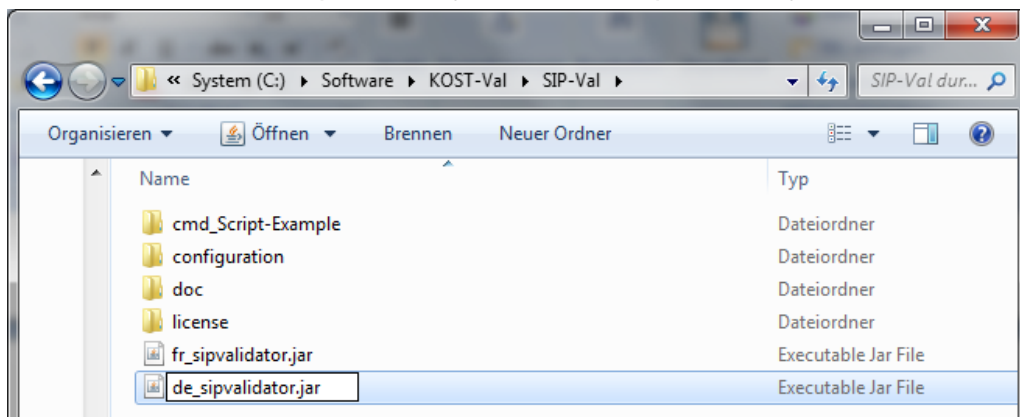
Ce script exécute en plus la validation de formats (3c) et la validation de période (3d). Ceci est également indiqué dans le résumé. En plus, le fichier "report.xml" est copié dans le dossier "logs" à la fin du script.

6.3 « commande_manuelle_SIP-Val.cmd »

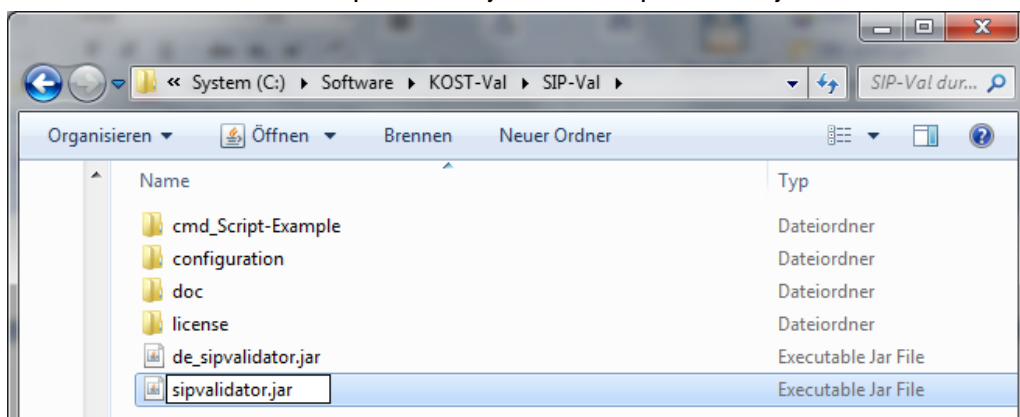
Ce script initialise les paramètres Java pour la fenêtre cmd courante sur l'application « Drittapplikationen\jre6\bin\java.exe ». En plus, un exemple de la saisie manuelle est indiqué qui peut servir de modèle.

7 Modifications pour la version française

- 1 Renommer le fichier « sipvalidator.jar » en « de_sipvalidator.jar ».



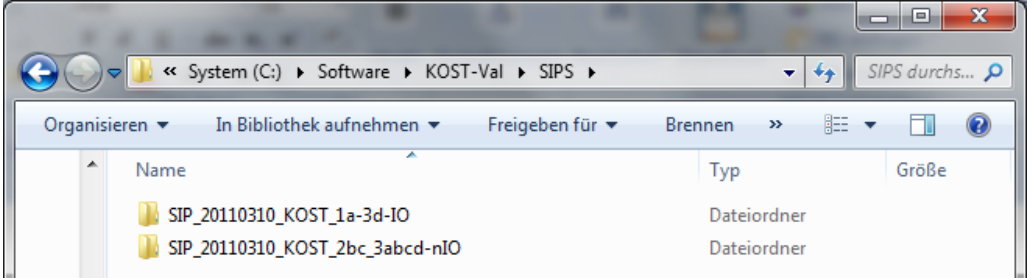
- 2 Renommer le fichier « fr_sipvalidator.jar » en « sipvalidator.jar ».

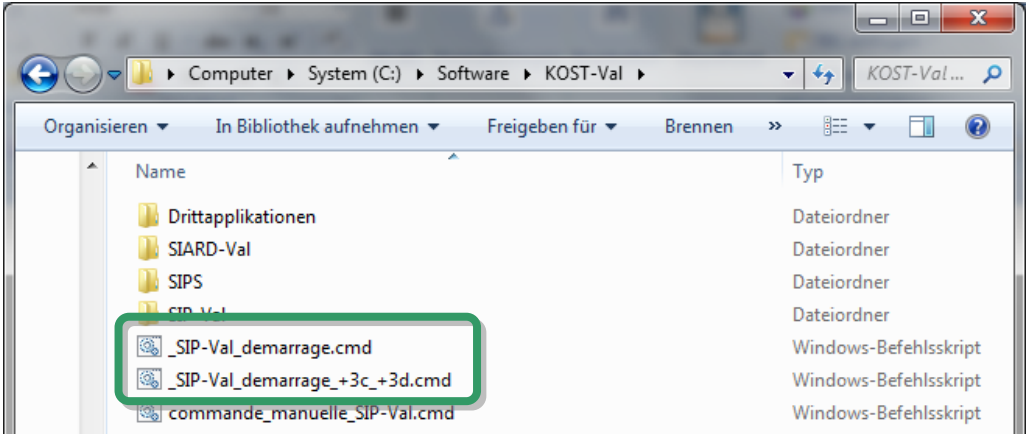


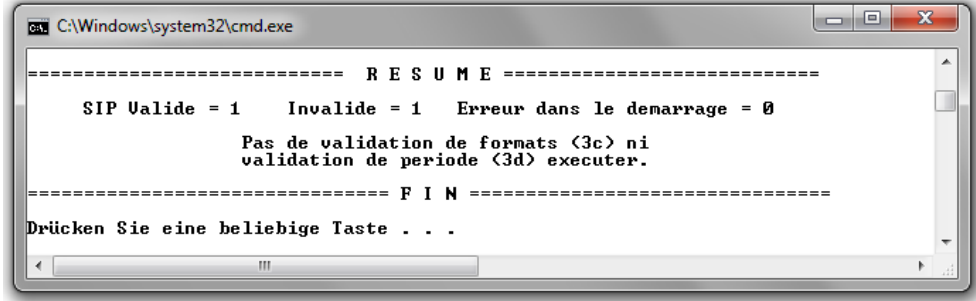
8 Exécuter la validation

8.1 Exécuter en utilisant le script de validation

- 1 Copier les SIPs à valider dans le dossier « KOST-Val\SIPS ».

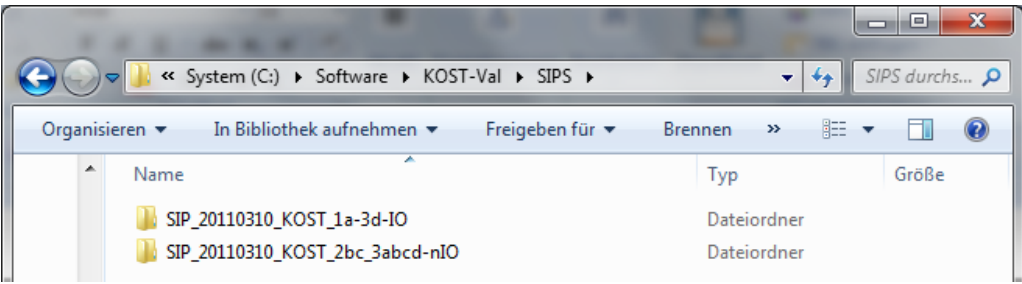
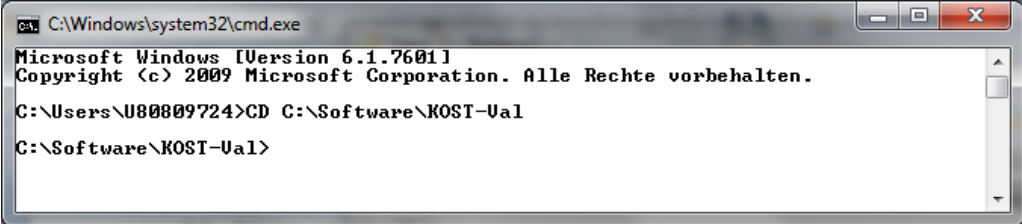
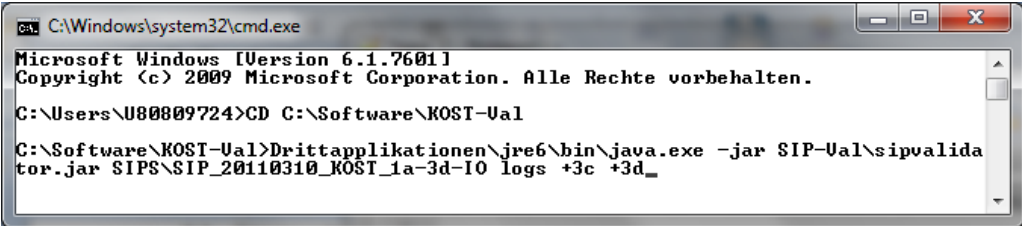

- 2 Lancer le script en double-cliquant sur « _SIP-Val_demarrage.cmd » ou « _SIP-Val_demarrage_+3c_+3d.cmd » (si la validation du format (3c) et la validation de période (3d) doivent aussi être effectuées) dans le dossier « KOST-Val ».


- 3 Tous les SIP ont été validés une fois que le résumé est affiché dans la fenêtre cmd.



Les résultats détaillés se trouvent dans le dossier « logs ».

8.2 Démarrer la validation manuellement

1	<p>Pour valider mettre en place d'abord le SIP.</p> 
2	<p>Ouvrir la fenêtre d'exécution (démarrer → Exécuter... → cmd [OK]) et aller dans le répertoire de travail (CD C:\Software\KOST-Val)⁷.</p> 
3	<p>Appeler le programme SIP-Val (en séparant tous les composants d'entrée par des espaces).</p>  <p>A Drittapplikationen\jre6\bin\java.exe -jar ou java -jar B SIP-Val\sipvalidator.jar C SIPS\SIP_20110310_KOST_1a-3d-IO D logs E +3c +3d</p> <p>A = appel de Java (Drittapplikationen\jre6\bin\java.exe -jar ou java -jar⁸), B = lien relationnel et fichier sipvalidator.jar, C = lien relationnel et nom complet du SIP à valider, D = lien relationnel et nom du répertoire logs, E = appel des validations facultatives.</p> <p><u>Notes:</u> Quand un élément d'entrée contient des espaces, celui-ci doit être inscrit entre guillemets.</p>

⁷ Le disque est par exemple changé avec CD /D.

⁸ Utiliser la commande `java -jar` si le Java Runtime Environment (JRE) version 6 est la version standard. Voir aussi le chapitre 3.2.3, Installation de Java (version 6 ou plus récente).

Il est également possible d'appeler SIP-Val depuis tout autre emplacement. Cependant, cela nécessite l'utilisation de liens absolus, non seulement dans l'appel du programme lui-même, mais aussi dans le fichier de configuration « sipvalidator.conf.xml ».

```

C:\Windows\system32\cmd.exe
Microsoft Windows [Version 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. Alle Rechte vorbehalten.

C:\Users\U80809724>C:\Software\KOST-Ual\Driftapplikationen\jre6\bin\java.exe -jar
C:\Software\KOST-Ual\SIP-Ual\sipvalidator.jar C:\Software\KOST-Ual\SIPS\SIP_20
110310_KOST_1a-3d-IO C:\Software\KOST-Ual\logs +3c +3d_
  
```

- A C:\Software\KOST-Ual\Driftapplikationen\jre6\bin\java.exe -jar / java -jar
- B C:\Software\KOST-Ual\SIP-Ual\sipvalidator.jar
- C C:\Software\KOST-Ual\SIPS\SIP_20110310_KOST_1a-3d-IO
- D C:\Software\KOST-Ual\logs
- E +3c +3d

Une alternative est la saisie manuelle en utilisant le script « commande_manuelle_SIP-Val.cmd ». Ce script permet d'appeler SIP-Val un peu plus facilement:

java.exe -jar SIP-Val\sipvalidator.jar SIPS\mon_SIP logs [+3c] [+3d]

```

C:\Windows\system32\cmd.exe

Avec ce script la commande pour demarrer SIP-Ual est plus simple:
java.exe -jar SIP-Ual\sipvalidator.jar SIPS\mon_SIP logs [+3c] [+3d]

Microsoft Windows [Version 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. Alle Rechte vorbehalten.

C:\Software\KOST-Ual>_
  
```

Ici, la seconde ligne peut être utilisée comme modèle⁹.

- 4 La validation est terminée dès que « Validation terminée. » ou « Validation interrompue. » est affiché dans la fenêtre cmd.

```

C:\Windows\system32\cmd.exe

[3c] = Valid
[3d] = Valid

TOTAL = Valid

jhove-Report = C:\Software\KOST-Ual\logs\SIP_20110310_KOST_1a-3d-IO.zip.jhove-lo
pdftron-Report = C:\Software\KOST-Ual\logs\SIP_20110310_KOST_1a-3d-IO.zip.pdftr
SIARD-Ual-Report = C:\Software\KOST-Ual\logs\SIP_20110310_KOST_1a-3d-IO.zip.siar

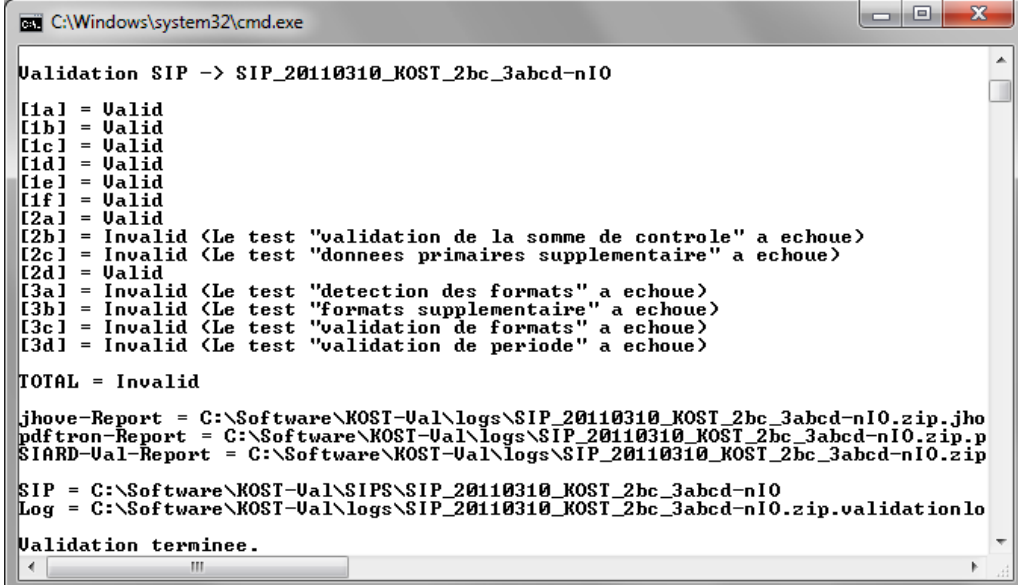
SIP = C:\Software\KOST-Ual\SIPS\SIP_20110310_KOST_1a-3d-IO
Log = C:\Software\KOST-Ual\logs\SIP_20110310_KOST_1a-3d-IO.zip.validationlog.log

Validation terminée.
  
```

Les résultats détaillés se trouvent dans le dossier « logs ».

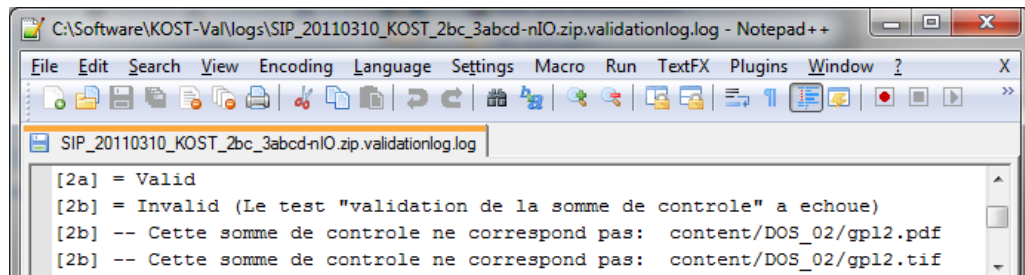
⁹ Utiliser la souris pour sélectionner le texte de la deuxième ligne, cliquer sur le bouton droit de la souris pour l'insérer dans la mémoire temporaire, insérer le texte dans la dernière ligne en cliquant à nouveau sur le bouton droit. Modifier ensuite la commande en effaçant les validations facultatives ou les crochets. Utiliser la touche flèche pour avancer à « mon_SIP », supprimer « mon_SIP » et entrer le nom du SIP. Pour finir, démarrer SIP-Val avec la touche Entrée.

9 Évaluer la validation

1	<p>Résumé des étapes de validation:</p> <table border="1"> <thead> <tr> <th>Étape de validation</th><th>Nom de l'étape (Stepname)</th></tr> </thead> <tbody> <tr><td>1a (abandon en cas d'erreur)</td><td>lisibilité</td></tr> <tr><td>1b (abandon en cas d'erreur)</td><td>validation de la structure des données</td></tr> <tr><td>1c (abandon en cas d'erreur)</td><td>nommage de la structure et des fichiers</td></tr> <tr><td>1d (abandon en cas d'erreur)</td><td>validation du schéma metadata.xml</td></tr> <tr><td>1e</td><td>déterminer le type du SIP"</td></tr> <tr><td>1f</td><td>existence des données primaires</td></tr> <tr><td>2a</td><td>données primaires manquant</td></tr> <tr><td>2b</td><td>validation de la somme de contrôle</td></tr> <tr><td>2c</td><td>données primaires supplémentaire</td></tr> <tr><td>2d</td><td>description des données primaires GEVER</td></tr> <tr><td>3a</td><td>détection des formats</td></tr> <tr><td>3b</td><td>formats supplémentaire</td></tr> <tr><td>3c (facultatif)</td><td>validation de formats</td></tr> <tr><td>3d (facultatif)</td><td>validation de période</td></tr> </tbody> </table>	Étape de validation	Nom de l'étape (Stepname)	1a (abandon en cas d'erreur)	lisibilité	1b (abandon en cas d'erreur)	validation de la structure des données	1c (abandon en cas d'erreur)	nommage de la structure et des fichiers	1d (abandon en cas d'erreur)	validation du schéma metadata.xml	1e	déterminer le type du SIP"	1f	existence des données primaires	2a	données primaires manquant	2b	validation de la somme de contrôle	2c	données primaires supplémentaire	2d	description des données primaires GEVER	3a	détection des formats	3b	formats supplémentaire	3c (facultatif)	validation de formats	3d (facultatif)	validation de période
Étape de validation	Nom de l'étape (Stepname)																														
1a (abandon en cas d'erreur)	lisibilité																														
1b (abandon en cas d'erreur)	validation de la structure des données																														
1c (abandon en cas d'erreur)	nommage de la structure et des fichiers																														
1d (abandon en cas d'erreur)	validation du schéma metadata.xml																														
1e	déterminer le type du SIP"																														
1f	existence des données primaires																														
2a	données primaires manquant																														
2b	validation de la somme de contrôle																														
2c	données primaires supplémentaire																														
2d	description des données primaires GEVER																														
3a	détection des formats																														
3b	formats supplémentaire																														
3c (facultatif)	validation de formats																														
3d (facultatif)	validation de période																														
2	<p>Après le démarrage de SIP-Val les résultats des étapes de validation sont affichés dans la console. Si l'étape de validation a été réalisée sans erreur, Valid est affiché, p.ex. [2a] = Valid.</p> <p>Si une étape de validation a échoué, Invalid et le nom de l'étape de validation sont affichés, p.ex. [2b] = Invalid (Le test "validation de la somme de controle" a echoue).</p> <p>A la fin, le résultat global et les liens vers les fichiers sont indiqués.</p>  <pre> C:\Windows\system32\cmd.exe Validation SIP -> SIP_20110310_KOST_2bc_3abcd-n10 [1a] = Valid [1b] = Valid [1c] = Valid [1d] = Valid [1e] = Valid [1f] = Valid [2a] = Valid [2b] = Invalid <Le test "validation de la somme de controle" a echoue> [2c] = Invalid <Le test "donnees primaires supplementaire" a echoue> [2d] = Valid [3a] = Invalid <Le test "detection des formats" a echoue> [3b] = Invalid <Le test "formats supplementaire" a echoue> [3c] = Invalid <Le test "validation de formats" a echoue> [3d] = Invalid <Le test "validation de periode" a echoue> TOTAL = Invalid jhove-Report = C:\Software\KOST-Ual\logs\SIP_20110310_KOST_2bc_3abcd-n10.zip.jho pdftron-Report = C:\Software\KOST-Ual\logs\SIP_20110310_KOST_2bc_3abcd-n10.zip.p SIARD-Val-Report = C:\Software\KOST-Ual\logs\SIP_20110310_KOST_2bc_3abcd-n10.zip SIP = C:\Software\KOST-Ual\SIPS\SIP_20110310_KOST_2bc_3abcd-n10 Log = C:\Software\KOST-Ual\logs\SIP_20110310_KOST_2bc_3abcd-n10.zip.validationlo Validation terminee. </pre> <p>Lors d'un appel via le script, un résumé de tous les SIPs est affiché.</p>																														

3

La sortie du log spécifie des informations détaillées supplémentaires sur les étapes de validation, en particulier sur les invalides. Sinon, elle est organisée comme la sortie de la console.

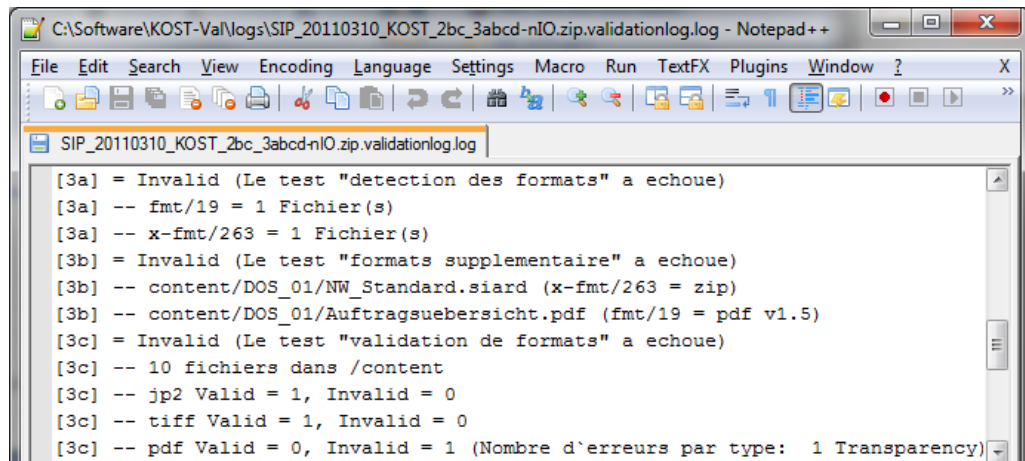


```

C:\Software\KOST-Va\Nogs\SIP_20110310_KOST_2bc_3abcd-nIO.zip.validationlog.log - Notepad++
File Edit Search View Encoding Language Settings Macro Run TextFX Plugins Window ? X
SIP_20110310_KOST_2bc_3abcd-nIO.zip.validationlog.log
[2a] = Valid
[2b] = Invalid (Le test "validation de la somme de controle" a echoue)
[2b] -- Cette somme de controle ne correspond pas: content/DOS_02/gpl2.pdf
[2b] -- Cette somme de controle ne correspond pas: content/DOS_02/gpl2.tif

```

Les analyses des applications tierces optionnelles employées sont enregistrées dans le répertoire respectif (voir aussi chapitre 5). Elles sont surtout établies pour les spécialistes puisqu'elles contiennent des indications techniques sur la validation de format. Pour la même raison, dans le cas de la sortie log de l'étape de validation 3c, les informations les plus compréhensibles ont été regroupées lorsque cela était possible.



```

C:\Software\KOST-Va\Nogs\SIP_20110310_KOST_2bc_3abcd-nIO.zip.validationlog.log - Notepad++
File Edit Search View Encoding Language Settings Macro Run TextFX Plugins Window ? X
SIP_20110310_KOST_2bc_3abcd-nIO.zip.validationlog.log
[3a] = Invalid (Le test "detection des formats" a echoue)
[3a] -- fmt/19 = 1 Fichier(s)
[3a] -- x-fmt/263 = 1 Fichier(s)
[3b] = Invalid (Le test "formats supplementaire" a echoue)
[3b] -- content/DOS_01/NW_Standard.siard (x-fmt/263 = zip)
[3b] -- content/DOS_01/Auftragsuebersicht.pdf (fmt/19 = pdf v1.5)
[3c] = Invalid (Le test "validation de formats" a echoue)
[3c] -- 10 fichiers dans /content
[3c] -- jpe2 Valid = 1, Invalid = 0
[3c] -- tiff Valid = 1, Invalid = 0
[3c] -- pdf Valid = 0, Invalid = 1 (Nombre d'erreurs par type: 1 Transparency)

```

Le SIP existant contient 10 fichiers primaires ([3c] -- 10 fichiers dans /content). Deux d'entre eux ne satisfont pas la reconnaissance de format ([3a] -- fmt/19 = 1 Fichier(s) et [3a] -- x-fmt/263 = 1 Fichier(s)). Dans l'étape 3b, il est indiqué: [3b] -- content/DOS_01/NW_Standard.siard (x-fmt/263 = zip) (ce fichier est au format ZIP) et [3b] -- content/DOS_01/Auftragsuebersicht.pdf (fmt/19 = pdf v1.5) (n'est pas un PDF/A mais seulement un PDF 1.5). Ceci signifie que 8 fichiers ont été reconnus comme formats conformes aux archives. Dans l'étape 3c, lorsque cela est possible, ces 8 fichiers sont vérifiés dans un validateur. Dans l'exemple il s'agit d'un JPEG2000 (valide), de 2 TIF (une fois valide et une fois invalide), d'un WAV (invalide) et d'un PDF/A (invalide). Les 3 fichiers restants n'ont pas pu être validés (fichiers CSV et TXT). Pour les fichiers SIARD et PDF/A invalides, le nombre d'erreurs par groupe principal est en plus indiqué (Invalid = 1 (Nombre d'erreurs par type: 1 Transparency))¹⁰. Ces indications doivent servir à ce que l'archiviste puisse librement déterminer la façon dont il doit poursuivre avec le SIP disponible et ses résultats.

¹⁰ Voir aussi les études du CECO sur les PDF/A, http://kost-ceco.ch/cms/index.php?studies_fr.

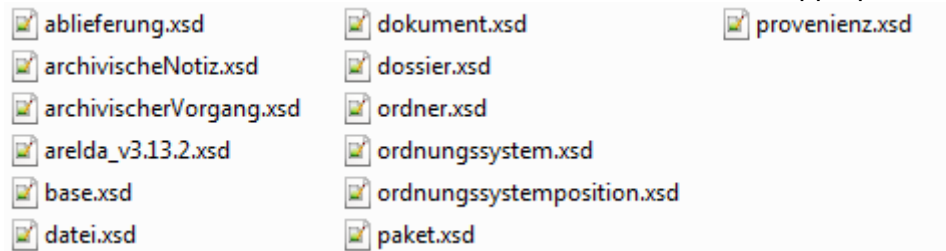
10 Description des modules de validation et des étapes de validation

Chaque étape de validation est décrite ci-après de façon détaillée; notamment quels critères ont été vérifiés et la manière dont ceux-ci sont apparus. Ces critères constituent un relevé des critères obligatoires de la spécification SIP v1.

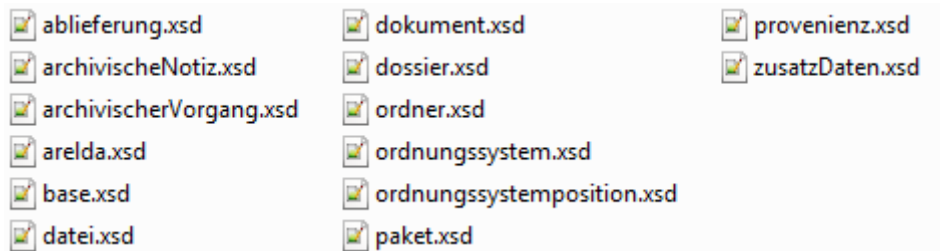
10.1 Module 1: Consistance paquet et XML

- a) Le SIP peut-il être ouvert sans erreur? Les conteneurs suivants doivent pouvoir être lus:
- Dossier (aucune terminaison supplémentaire)
 - ZIP
 - ZIP64
- b) Existe t-il une structure de répertoire primaire correcte? Le SIP est valide si une des trois variantes de structures suivantes est exacte (défini respectivement par deux chemins):
- [SIP-Name]/header/metadata.xml et [SIP-Name]/content
 - [SIP-Name].zip/header/metadata.xml et [SIP-Name].zip/content
 - [SIP-Name].zip/[SIP-Name]/header/metadata.xml et [SIP-Name].zip/[SIP-Name]/content
- Remarque: [Nom-SIP] doit toujours être identique à l'intérieur de la structure du répertoire (concernant la variante 3).
- c) Les noms de répertoire et de fichier correspondent-ils aux restrictions de la spécification?
- Critères:
- Police de caractères: Les spécifications concernant la police de caractères à utiliser pour la dénomination des dossiers et des fichiers doivent être entièrement respectées. La police de caractères autorisée pour la dénomination des dossiers et des fichiers est un sous-ensemble de l'US-ASCII (les codes de caractères des caractères autorisés sont identiques à l'US-ASCII, l'ISO-8859 et l'UTF-8 dans les polices de caractères). Les caractères suivants sont autorisés:
 - Lettres: A-Z a-z
 - Chiffres: 0-9
 - Autres: ! # \$ % () + , - _ . = @ [] { } ~ ESPACE
 - Nom du dossier de niveau supérieur: Dans le cas d'un SIP, le nom du dossier de niveau supérieur commence toujours par la suite de caractères SIP_ suivie des indications définies à identifier. Le nom du dossier de niveau supérieur est toujours établi selon le modèle suivant:
SIP_[date de versement]_[nom du service versant](_[référence]).
La restriction de dénomination est enregistrée dans le fichier de configuration, voir aussi 5.1.3 <allowsipname>.
 - Nom du dossier de second niveau: Le dossier de niveau supérieur contient toujours deux dossiers. Ces dossiers portent les noms de header/ et content/. Il est interdit de placer d'autres dossiers ou fichiers directement dans le dossier de niveau supérieur.
 - Dossier header/: Le dossier prédéfini dans le dossier header/ d'un SIP pour le versement se nomme: xsd/. La présence d'autres dossiers dans le dossier header/ n'est pas autorisée. Le fichier prédéfini dans le dossier header/ se nomme toujours metadata.xml. Ce fichier est le seul à être autorisé dans le dossier header/.

- Nom des fichiers xsd: Les fichiers suivants doivent pour la version 1 du AFS-SIP être inclus dans header/xsd/ avec la dénomination appropriée.



Pour la version 4 du AFS-SIP respectivement eCH-0160 version 1 se sont les fichiers suivants:



Aucun autre fichier n'est autorisé.

- Selon le fichier arelda dans le SIP il s'agit de la version 1 ou 4. Ceci est validé les versions souhaitées dans le fichier de configuration. Les Versions non autorisées sont émis.
 - Longueur des chemins: La longueur du chemin pour chaque fichier et pour chaque dossier à l'intérieur des paquets d'informations ne doit pas dépasser 180 caractères. Le chemin contient d'ailleurs toujours le dossier de niveau supérieur¹¹. Les barres obliques (/) doivent également être comptées. Le nombre de caractères autorisé est enregistré dans le fichier de configuration.
 - Longueur des dossiers ou des fichiers (uniquement pour la version 1): Le nom d'un dossier ou d'un fichier ne doit pas dépasser 40 caractères au maximum. L'extension du fichier (File Extension) ainsi que le point entre la désignation et l'extension doivent être comptés¹². Le nombre de caractères autorisé est enregistré dans le fichier de configuration.
- d) Les métadonnées .xml correspondent-elles aux fichiers schéma dans header/xsd? Les éventuelles erreurs XML sont affichées¹³.
- e) Le type SIP est transmis et affiché (metadata.xml)//ablieferungstyp: GEVER ou FILE
- f) Il est indiqué si aucun fichier primaire ne se trouve dans le répertoire content/ (GEVER-SIP correct uniquement pour l'archivage d'une structure de dossier, mais FILE-SIP manquant).

¹¹ Concernant le dossier de niveau supérieur, le nombre de caractères est toujours inclus avec l'extension .zip, même si le SIP a été transmis en tant que structure de dossier. Ceci correspond parfois à 4 caractères supplémentaires dans la longueur de chemin.

¹² Concernant le dossier de niveau supérieur, le nombre de caractères est toujours inclus avec l'extension .zip, même si le SIP a été transmis en tant que structure de dossier. Ceci correspond parfois à 4 caractères supplémentaires dans la longueur du fichier ou du dossier.

¹³ Le schéma arelda_v3.13.2.xsd contient une erreur en ce qui concerne l'algorithme et la somme de vérification (lignes 924 et 925). Ces éléments doivent apparaître et devraient contenir l'entrée correspondante minOccurs="1". Cette erreur sera éliminée par les AFS. Jusqu'à présent le SIP-Val ne fait aucune objection lorsque ces éléments manquent, même s'ils devraient être présents selon la spécification v1 du SIP.

10.2 Module 2: Consistance du fichier

10.2.1 Les étapes de validation obligatoires

- a) Il faut vérifier si tous les fichiers référencés existent dans (metadata.xml) /paket/inhaltsverzeichnis. Lister les fichiers manquants.
- b) Valider les fichiers avec somme de contrôle dans (metadata.xml) // pruefsumme. Les fonctions de hachage suivants sont autorisés: MD5, SHA-1, SHA-256 et SHA-512.
- c) Lister les fichiers et répertoires enregistrés dans le SIP existant, mais pas dans le (metadata.xml) /paket/inhaltsverzeichnis (sauf le fichier metadata.xml).
- d) Concernant GEVER-SIP, il faut vérifier si tous les fichiers référencés dans (metadata.xml) /paket/inhaltsverzeichnis/content sont aussi enregistrés dans (metadata.xml) /paket/ablieferung/ordnungssystem. Lister les éventuelles inconsistances. (//dokument[@id] => //datei[@id]).

10.3 Module 3: Consistance du format de fichier et de la date

10.3.1 Les étapes de validation obligatoires

- a) Effectuer une reconnaissance de format avec DROID (et le File format registry PRONOM) et comparer les formats avec la liste des formats autorisés. Cette liste est enregistrée dans le fichier de configuration, voir aussi 5.1.12 <allowedformats>.
- b) Lister tous les fichiers non spécifiés avec les indications de format correspondantes.

10.3.2 Les étapes de validation optionnelles

- c) (*commutable*) Validation de format, avec JHOVE ainsi qu'avec des validateur externe (par ex. SIARD: SIARD-Val [KOST] et PDF/A-1b: PDF/A Manager [PDFTRON]) de tous les fichiers sélectionnés en fonction du format dans /content (liste à configurer des formats de fichiers à valider, voir aussi 5.1.13 <validatedformats>).
- d) (*commutable*) Validation de période: Vérification des indications temporelles dans (metadata.xml)/versement.

Les éléments suivants mentionnés de manière hiérarchique sont vérifiés:

Élément	Indications temporelles
Versement	Période de création (à partir de - jusqu'à)
Dossier ¹⁴	Période de création (à partir de - jusqu'à)
Document	Période de création (à partir de - jusqu'à)

Si un élément transgresse au moins un des critères suivants (et l'étape de validation de la période), il est classé comme invalide:

- La valeur temporelle « jusqu'à » doit être supérieure ou égale à la valeur temporelle « à partir de » et aucune des deux valeurs ne doit se trouver dans le futur.
- Les indications temporelles d'un élément doivent se situer pendant la période de création de l'élément supérieur.

¹⁴ Les périodes de création doivent être spécifiées uniquement dans le dossier. Tous les types d'éléments restants sont aussi valides, si aucune période de création n'est mentionnée.

11 Droits d'auteur

SIP-Val a été développé par le CECO. Le CECO en possède tous les droits. SIP-Val a été publié par le CECO en 2011 sous la licence GNU General Public License v3+.

Notice:	This product includes software developed by the Apache Software Foundation (http://www.apache.org/).
----------------	---

SIP-Val se base sur les composants inchangés suivants d'autres fabricants qui sont intégrés directement dans le code source de SIP-Val:

Programme tiers/Composants tiers	Version	Licence
Apache Commons http://commons.apache.org/ - commons-collections-3.2.1.jar - commons-configuration-1.6.jar - commons-digester-1.8.jar - commons-lang-2.4.jar - commons-logging-1.1.1.jar	3.2.1 1.6 1.8 2.4 1.1.1	Apache License 2.0
Apache log4j http://logging.apache.org/log4j/	1.2.12	Apache License 2.0
Apache Xalan-Java http://xml.apache.org/xalan-j/	2.7.0	Apache License 2.0
Apache Xerces http://xerces.apache.org/	2.7.1	Apache License 2.0
DROID http://sourceforge.net/apps/mediawiki/droid/	5.0.3	3c BSD- License
Junit 4.4 http://www.junit.org/	4.4	CPL v1.0
Spring Framework API http://static.springsource.org/spring/docs/3.0.x/api/	3.0.0	Apache License 2.0
zip64 http://sourceforge.net/projects/zip64file/	1.02	GPL v2+ License

SIP-Val se base sur la bibliothèque inchangée suivante d'autres fabricants qui doit être installée séparément par l'utilisateur, afin que SIP-Val fonctionne:

Programme tiers/Composants tiers	Version	Licence
JHOVE http://hul.harvard.edu/jhove/	1.5	LGPL v2.1 License

Les utilisateurs de SIP-Val sont tenus de suivre les dispositions de licence de tous ces composants qui se trouvent dans le répertoire SIP-Val\license.

Remarque

Pour l'exécution d'une validation PDF/A (étape de validation optionnelle 3c), l'application indirecte de la version test du logiciel commercial PDF/A Manager de PDFTRON a été mise en œuvre (au moyen de l'appel de la console) et pour la validation SIARD le CECO Tools SIARD-Val. Ces logiciels ne sont pas un composant de SIP-Val et doivent être fournis séparément et installés pour une éventuelle utilisation. Des informations supplémentaires sont disponibles au chapitre 4.2.1.