

Aide pour le logiciel Vitam Tool

Version 2.0.2

Table des matières

Installation.....	1
Lancement.....	2
Points d'attention ou Limitations.....	3
Mémoire.....	3
Fichier.....	3
LibreOffice/OpenOffice.....	3
Extraction des mots clefs.....	3
Interface Ligne de commande.....	4
Quelques exemples d'utilisation.....	5
Analyse d'un fichier XML.....	5
Analyse d'un document ou d'un répertoire.....	6
Calcul d'empreintes.....	7
Encodage en PDF/A-1B.....	8
Quelques exemples.....	9
Interface graphique.....	11
Affichage principal.....	11
Subdivision de l'écran.....	13

Installation

Prérequis :

- une JVM (ou mieux une JDK/JSE) installée en version 1.6 minimum,
- au moins 1Go de mémoire
- au moins 1 Go d'espace disque disponible

Un zip ou 7z est fourni est contient l'ensemble des binaires. Installez ce répertoire où vous le souhaitez.

Un fichier de configuration XML (vitam-config.xml) permet de changer les paramètres par défaut.

En ouvrant l'interface graphique, il est possible de changer les paramètres et de les sauver.

Lancement

Deux modes existent :

1. Ligne de commande
 1. Sous Linux : depuis le répertoire de l'application
`java -cp VitamToolGui.jar fr.gouv.culture.vitam.command.VitamCommand`
 2. Sous Windows : depuis le répertoire de l'application
Idem ou `VitamToolCmd.exe`
2. Avec interface graphique
 1. Sous Linux :
Double cliquez sur `VitamToolGui.jar` ou `java -cp VitamToolGui.jar`
 2. Sous Windows :
Idem ou double cliquez sur `VitamToolGui.exe`

Points d'attention ou Limitations

Mémoire

Ce logiciel consomme de la mémoire. Le PC/Serveur hébergeant ce logiciel doit faire attention à la disponibilité de la mémoire. Via les commandes exécutables, le logiciel est configuré pour utiliser les paramètres -Xms200m et -Xmx1024m.

Vous pouvez modifier ces paramètres en utilisant explicitement le lancement via java et non via l'exécutable :

`-Xms200m -Xmx1024m`

Fichier

Les fichiers générés seront au format UTF-8. Pour assurer le bon fonctionnement, si vous lancer via « java », veuillez à utiliser les options suivantes :

`-Dfile.encoding=UTF-8 -Dclient.encoding.override=UTF-8`

LibreOffice/OpenOffice

Pour les conversions en PDF/A-1B, LibreOffice/OpenOffice est utilisé en mode batch.

Attention, si une instance tourne en dehors de ce logiciel, elle sera arrêtée brutalement.

A noter également que lors de la conversion, le répertoire cible ne DOIT PAS contenir d'espace dans son nom.

Extraction des mots clefs

Ce processus peut être consommateur en mémoire (lecture intégrale du document et gestion des occurrences). Aussi, veuillez à limiter son usage sur des fichiers de tailles raisonnables ou à allouer suffisamment de mémoire au programme.

Interface Ligne de commande

Sans argument, la commande affichera l'aide sommaire suivante :

```
Necessite au moins "--xml fichier" ou "--print fichier" ou "--convertpdfa source
destination" ou "--checkdigest fichier" ou "--createdigest 5-args" comme argument
    Obligatoirement un parmi (-m,--xml fichier | -p,--print fichier | -3,--
convertpdfa source destination | -4,--checkdigest fichier | -5,--createdigest 5-args)
    ou -5,--createdigest source cible (-notar | fichiertar) (-noglobal |
globaldir/prefix) (-noperfile | -perfile) [5 arguments obligatoires]
    [-d,--xsd fichier (default=/resources/seda/v10/seda_v1-0_archivetransfer.xsd) |
-nd,--notxsd pour ne pas avoir de check XSD]
    [-r,--root racine du xsd (default=ArchiveTransfer)]
    [-ff,--filefield champ (default=Attachment)]
    [-fa,--fileattrib attribut (default=@filename)]
    [-ma,--mimeattrib attribut (default=@mimeType)]
    [-ta,--formatattrib attribut (default=@format)]
    [-df,--digestfield champ (default=Integrity)]
    [-aa,--algoattrib attribut (default=@algorithme)]
    [-s,--schematron fichier (default=/resources/seda/seda_v10.sch) | -ns,--
notschematron pour ne pas avoir d action schematron]
    [-l,--xsl fichier (default=/resources/seda/seda2ead-1.0.xsl) | -nl,--notxsl pour
ne pas avoir d action XSL transformation]
    [-g,--rng fichier (non fait par default)]
    [-n,--signature fichier (default=droid-6.1/DRUID_SignatureFile_V63.xml)]
    [-c,--container fichier (default=droid-6.1/container-signature-20120828.xml)]
    [-f,--checkformat (default=faux)]
    [-a,--checkarchives (default=false) | -na,--notcheckarchives]
    [-v,--checkrecursive (default=false) | -nv,--notcheckrecursive]
    [-h,--computesha algo (où algo=SHA-1,SHA-256,SHA-512 ou un sous-ensemble,
default=true:false:false) | -nh,--notcomputesha]
    [-e,--extensionrecur filtre_liste_a_virgule (default=pas de filtre sur
l'extension)]
    [-o,--showformat (default=faux)]
    [-x,--formatoutput format (in TXT|XML|XMLS, default=XML)]
    [-1,--outputfile fichier (default=STDOUT)]
    [-2,--outputarchive xslfichier nom_de_base_fichier (default=none)]
    [-z,--preventxfmt (default=true) | -nz,--notpreventxfmt]
    [-3,--convertpdfa source destination (default=faux)]
    [-k,--msperkb msParKo (default=300)]
    [-t,--lowlimitms limiteBasseEnMs (default=60000)]
    [-y,--extractkeyword (default=true) | -ny,--notextractkeyword]
    [-i,--ranklimit limitRankOccur (default=10)]
    [-w,--wordlimit limitWordOccur (default=100)]
    [--help] pour imprimer cette aide
    [-0,--config configurationFile (default=vitam-config.xml)]
```

Copyright (c) 2012 Ministere de la Culture et de la Communication
Sous-Direction du Systeme d'Information
Projet Vitam

Version: 2.0.0

Contributeurs: Frederic Bregier

Site web: <http://www.archivesnationales.culture.gouv.fr/>

Licence: GPLv3

Quelques exemples d'utilisation

Dans ce qui suit, la commande de lancement sera VitamToolCmd.exe, mais peut être remplacée par les autres méthodes d'appel.

Cinq grandes fonctions sont possibles :

```
(-m,--xml fichier | -p,--print fichier | -3,--convertpdfa source destination | -4,--  
checkdigest fichier | -5,--createdigest 5-args)
```

Analyse d'un fichier XML

Avec l'option -m fichier.xml (analyse d'un fichier XML), les options suivantes sont prises en compte :

```
[-d,--xsd fichier (default=/resources/seda/v10/seda_v1-0_archivetransfer.xsd) |  
-nd,--notxsd pour ne pas avoir de check XSD]  
[-r,--root racine du xsd (default=ArchiveTransfer)]  
[-ff,--filefield champ (default=Attachment)]  
[-fa,--fileattrib attribut (default=@filename)]  
[-ma,--mimeattrib attribut (default=@mimeType)]  
[-ta,--formatattrib attribut (default=@format)]  
[-df,--digestfield champ (default=Integrity)]  
[-aa,--algoattrib attribut (default=@algorithme)]  
[-s,--schematron fichier (default=/resources/seda/seda_v10.sch) | -ns,--  
notschematron pour ne pas avoir d action schematron]  
[-l,--xsl fichier (default=/resources/seda/seda2ead-1.0.xsl) | -nl,--notxsl pour  
ne pas avoir d action XSL transformation]  
[-g,--rng fichier (non fait par default)]  
[-n,--signature fichier (default=droid-6.1/DRUID_SignatureFile_V63.xml)]  
[-c,--container fichier (default=droid-6.1/container-signature-20120828.xml)]  
[-f,--checkformat (default=faux)]  
[-a,--checkarchives (default=false) | -na,--notcheckarchives]  
[-h,--computesha algo (où algo=SHA-1,SHA-256,SHA-512 ou un sous-ensemble,  
default=true:false:false) | -nh,--notcomputesha]  
[-o,--showformat (default=faux)]  
[-x,--formatoutput format (in TXT|XML|XMLS, default=XML)]  
[-l,--outputfile fichier (default=STDOUT)]  
[-z,--preventxfmt (default=true) | -nz,--notpreventxfmt]  
[-y,--extractkeyword (default=true) | -ny,--notextractkeyword]  
[-i,--ranklimit limitRankOccur (default=10)]  
[-w,--wordlimit limitWordOccur (default=100)]  
[--help] pour imprimer cette aide  
[-0,--config configurationFile (default=vitam-config.xml)]
```

Analyse d'un document ou d'un répertoire

Avec l'option -p chemin (analyse d'un répertoire ou d'un document), les options suivantes sont prises en compte :

```
[-n,--signature fichier (default=droid-6.1/DROID_SignatureFile_V63.xml)]
[-c,--container fichier (default=droid-6.1/container-signature-20120828.xml)]
[-f,--checkformat (default=faux)]
[-a,--checkarchives (default=false) | -na,--notcheckarchives]
[-v,--checkrecursive (default=false) | -nv,--notcheckrecursive]
[-h,--computesha algo (où algo=SHA-1,SHA-256,SHA-512 ou un sous-ensemble,
default=true:false:false) | -nh,--notcomputesha]
[-e,--extensionrecur filtre_liste_a_virgule (default=pas de filtre sur
l'extension)]
[-o,--showformat (default=faux)]
[-x,--formatoutput format (in TXT|XML|XMLS, default=XML)]
[-1,--outputfile fichier (default=STDOUT)]
[-2,--outputarchive xslfichier nom_de_base_fichier (default=none)]
[-z,--preventxfmt (default=true) | -nz,--notpreventxfmt]
[-y,--extractkeyword (default=true) | -ny,--notextractkeyword]
[-i,--ranklimit limitRankOccur (default=10)]
[-w,--wordlimit limitWordOccur (default=100)]
[--help] pour imprimer cette aide
[-0,--config configurationFile (default=vitam-config.xml)]
```

Calcul d'empreintes

Avec l'option -4 chemin (calcul des empreintes d'un répertoire ou d'un document), les options suivantes sont prises en compte :

```
[-ff,--filefield champ (default=Attachment)]
[-fa,--fileattrib attribut (default=@filename)]
[-df,--digestfield champ (default=Integrity)]
[-aa,--algoattrib attribut (default=@algorithme)]
[-v,--checkrecursive (default=false) | -nv,--notcheckrecursive]
[-h,--computesha algo (où algo=SHA-1,SHA-256,SHA-512 ou un sous-ensemble,
default=true:false:false) | -nh,--notcomputesha]
[-e,--extensionrecur filtre_liste_a_virgule (default=pas de filtre sur
l'extension)]
[-o,--showformat (default=faux)]
[-x,--formatoutput format (in TXT|XML|XMLS, default=XML)]
[-l,--outputfile fichier (default=STDOUT)]
[--help] pour imprimer cette aide
[-0,--config configurationFile (default=vitam-config.xml)]
```

Calcul d'empreintes unitaires par lot

Avec l'option -5 5-arguments (calcul des empreintes avec de multiples options OBLIGATOIRES), les options suivantes sont prises en compte :

```
-5,--createdigest source cible (-notar | fichiertar) (-noglobal | globaldir/prefix) (-noperfile | -perfile) [5 arguments obligatoires]
    [-ff,--filefield champ (default=Attachment)]
    [-fa,--fileattrib attribut (default=@filename)]
    [-df,--digestfield champ (default=Integrity)]
    [-aa,--algoattrib attribut (default=@algorithme)]
    [-v,--checkrecursive (default=false) | -nv,--notcheckrecursive]
    [-h,--computesha algo (où algo=SHA-1,SHA-256,SHA-512 ou un sous-ensemble,
default=true:false:false) | -nh,--notcomputesha]
    [-e,--extensionrecur filtre_liste_a_virgule (default=pas de filtre sur
l'extension)]
    [--help] pour imprimer cette aide
    [-0,--config configurationFile (default=vitam-config.xml)]
```


Encodage en PDF/A-1B

Attention : toute instance de LibreOffice ou OpenOffice sera arrêtée.

Avec l'option -2 cheminSource cheminDestination (transformation en PDF/A-1B en respectant l'arborescence), les options suivantes sont prises en compte :

```
[-n,--signature fichier (default=droid-6.1/DROID_SignatureFile_V63.xml)]
[-c,--container fichier (default=droid-6.1/container-signature-20120828.xml)]
[-e,--extensionrecur filtre_liste_a_virgule (default=pas de filtre sur
l'extension)]
[-x,--formatoutput format (in TXT|XML|XMLS, default=XML)]
[-l,--outputfile fichier (default=STDOUT)]
[-z,--preventxfmt (default=true) | -nz,--notpreventxfmt]
[-3,--convertpdfa source destination (default=faux)]
[-k,--msperkb msParKo (default=300)]
[-t,--lowlimitms limiteBasseEnMs (default=60000)]
[--help] pour imprimer cette aide
[-0,--config configurationFile (default=vitam-config.xml)]
```

Quelques exemples

- Tester un fichier ArchiveTransfer conforme au XSD et au Schématron par défaut, et en appliquant une transformation vers EAD (incomplète)
`VitamToolCmd.exe -m archivetransfer.xml`
 - Tester un fichier ArchiveTransfer en changeant le XSD, le Schématron et la feuille de style XSLT
`VitamToolCmd.exe -m archivetransfer.xml -d fichier.xsd -s fichier.sch -l fichier.xsl`
 - Tester un fichier ArchiveTransfer en changeant le XSD mais en ne vérifiant pas le schématron ni la transformation XSLT
`VitamToolCmd.exe -m archivetransfer.xml -d fichier.xsd -ns -nl`
 - Tester un fichier ArchiveTransfer en changeant le XSD, en ne vérifiant pas le schématron ni la transformation XSLT et en changeant les champs et attributs utilisés
`VitamToolCmd.exe -m archivetransfer.xml -d fichier.xsd -ns -nl -r
ArchiveTransfer -ff Attachment -fa @filename -ma @mimecode -ta @format -df
Integrity -aa @algorithme`
 - Tester un fichier ArchiveTransfer sur un profil RNG (mais pas les autres aspects)
`VitamToolCmd.exe -m archivetransfer.xml -nd -ns -nl -g fichier.rng`
 - Tester un fichier ArchiveTransfer en vérifiant ses informations vis à vis des documents référencés
`VitamToolCmd.exe -m archivetransfer.xml -f`
 - Tester un fichier ArchiveTransfer en vérifiant ses informations vis à vis des documents référencés en spécifiant une autre base de signature Pronom
`VitamToolCmd.exe -m archivetransfer.xml -f -n DROID_SignatureFile_VXX.xml
-c container-signature-date.xml`
 - Extraire des données depuis un fichier ArchiveTransfer
`VitamToolCmd.exe -m archivetransfer.xml -o`
 - Extraire des données depuis un fichier ArchiveTransfer, y compris sur des archives (zip, tar, ...)
`VitamToolCmd.exe -m archivetransfer.xml -o -a`
 - Extraire des données depuis un fichier ArchiveTransfer en excluant les formats x-fmt (warning)
`VitamToolCmd.exe -m archivetransfer.xml -o -z`
 - Extraire des données depuis un fichier ArchiveTransfer, y compris les empreintes dans d'autres formats
`VitamToolCmd.exe -m archivetransfer.xml -o -h SHA-1,SHA-256,SHA-512`
 - Extraire des données depuis un fichier ArchiveTransfer en changeant le format de sortie
`VitamToolCmd.exe -m archivetransfer.xml -o -x XML`
 - Extraire des données depuis un fichier ou un répertoire
`VitamToolCmd.exe -p chemin -o`
 - Extraire des données depuis un fichier ou un répertoire récursivement et en testant les archives et en limitant à certaines extensions
`VitamToolCmd.exe -p chemin -o -v -a -e doc,docx,odt`
 - Extraire des données depuis un fichier ou un répertoire récursivement et en testant les archives et en limitant à certaines extensions, en enregistrant le résultat dans un fichier
`VitamToolCmd.exe -p chemin -o -v -a -e doc,docx,odt -l output.xml`
- Attention : Si plusieurs sorties sont demandées (-o -f par exemple), elles seront cumulatives et ne seront donc pas conforme à un XML. De même, si le mode est XMLS (un XML par fichier), le résultat sera la concaténation de ces fichiers, donc non conforme à un seul XML.

- Extraire des données depuis un fichier ou un répertoire récursivement en limitant les mots clefs en sortie aux mots occurrence au moins 10 et environ les 100 premiers.

```
VitamToolCmd.exe -p chemin -o -v -i 10 -w 100
```

- Extraire des données depuis un fichier ou un répertoire toutes options actives et le traduire avec une feuille de style XSLT (ici pour produire un ArchiveTransfer à compléter)

```
VitamToolCmd.exe -p chemin -o -v -a -i 1 -w 0 -h SHA-1,SHA-256,SHA-512 -x XML -1 outputSimple.xml -2 DROID2SEDA_filetype.xsl -z
```

- Transformer un dossier ou un fichier récursivement en PDF/A-1B en autorisant 300 ms en moyenne par KB et avec au moins 60 secondes de latence

```
VitamToolCmd.exe -3 cheminSource cheminDestination -v -k 300 -t 60000
```

- Calculer des empreintes un dossier ou un fichier récursivement

```
VitamToolCmd.exe -4 cheminSource -v
```

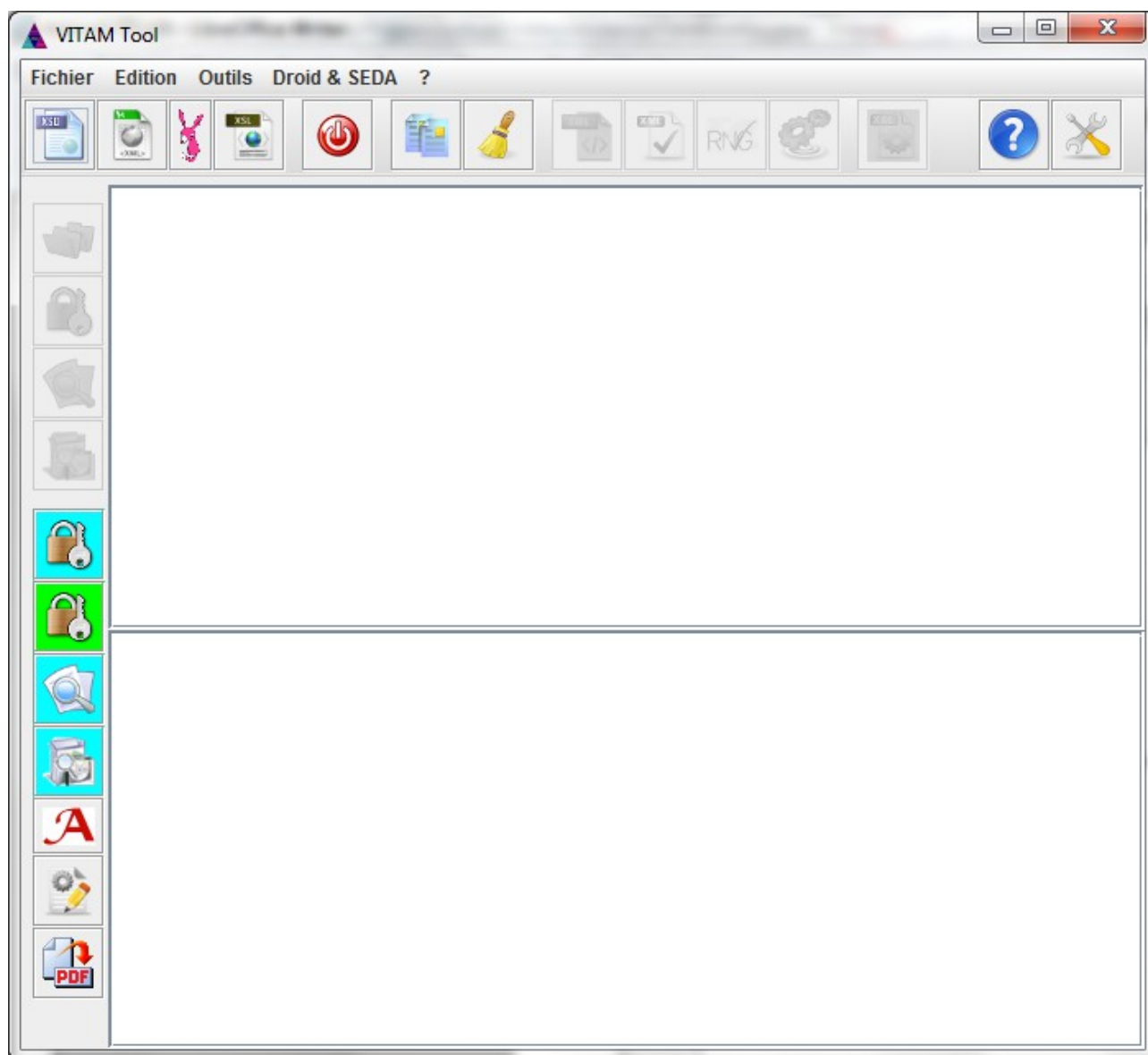
- Calculer des empreintes un dossier ou un fichier récursivement avec multiples sorties

```
VitamToolCmd.exe -5 cheminSource cheminCible -notar chemincible/global -perfile -v
```

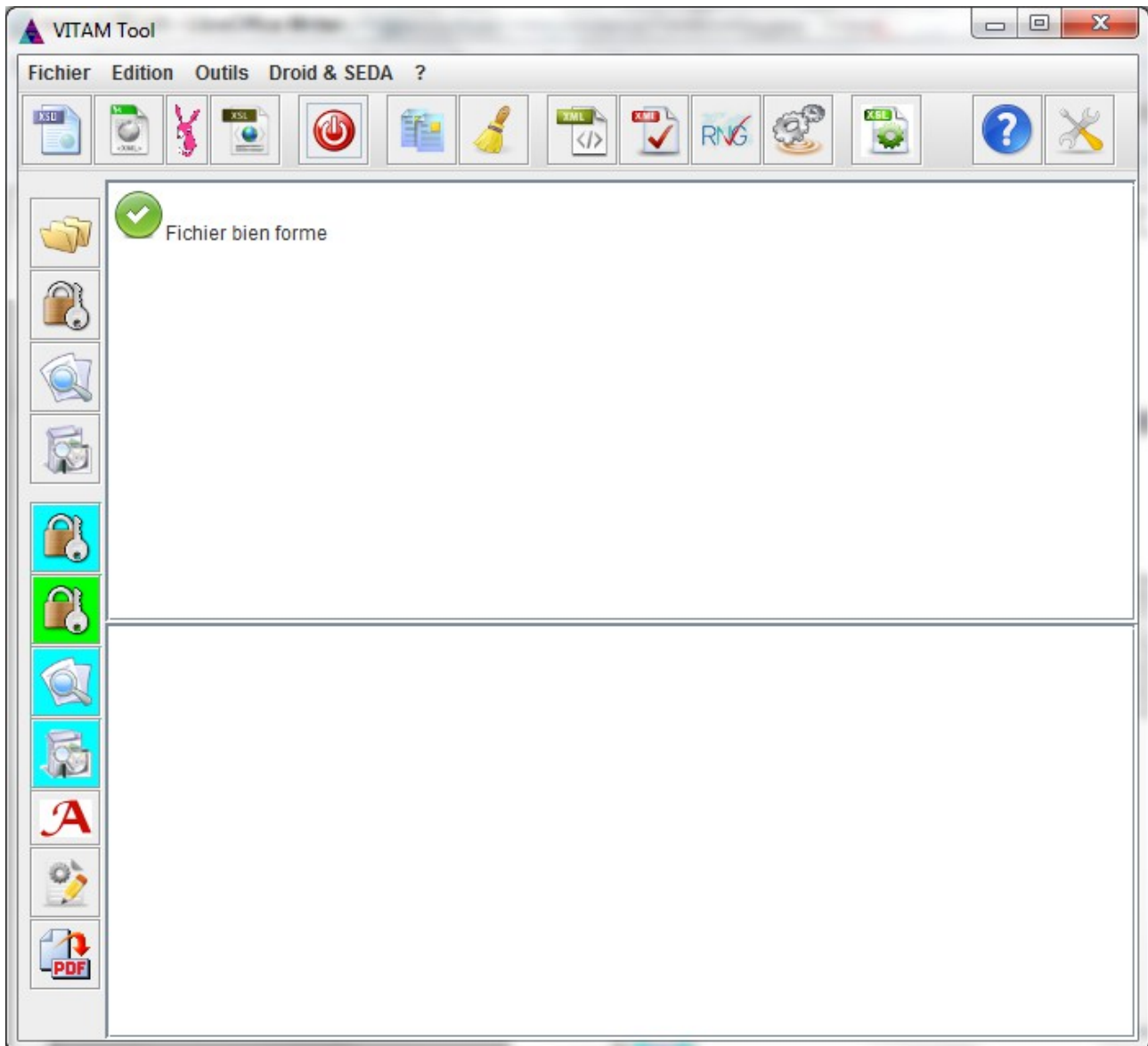
Interface graphique

Affichage principal

Au démarrage, sans qu'aucun fichier XML n'ait été chargé.



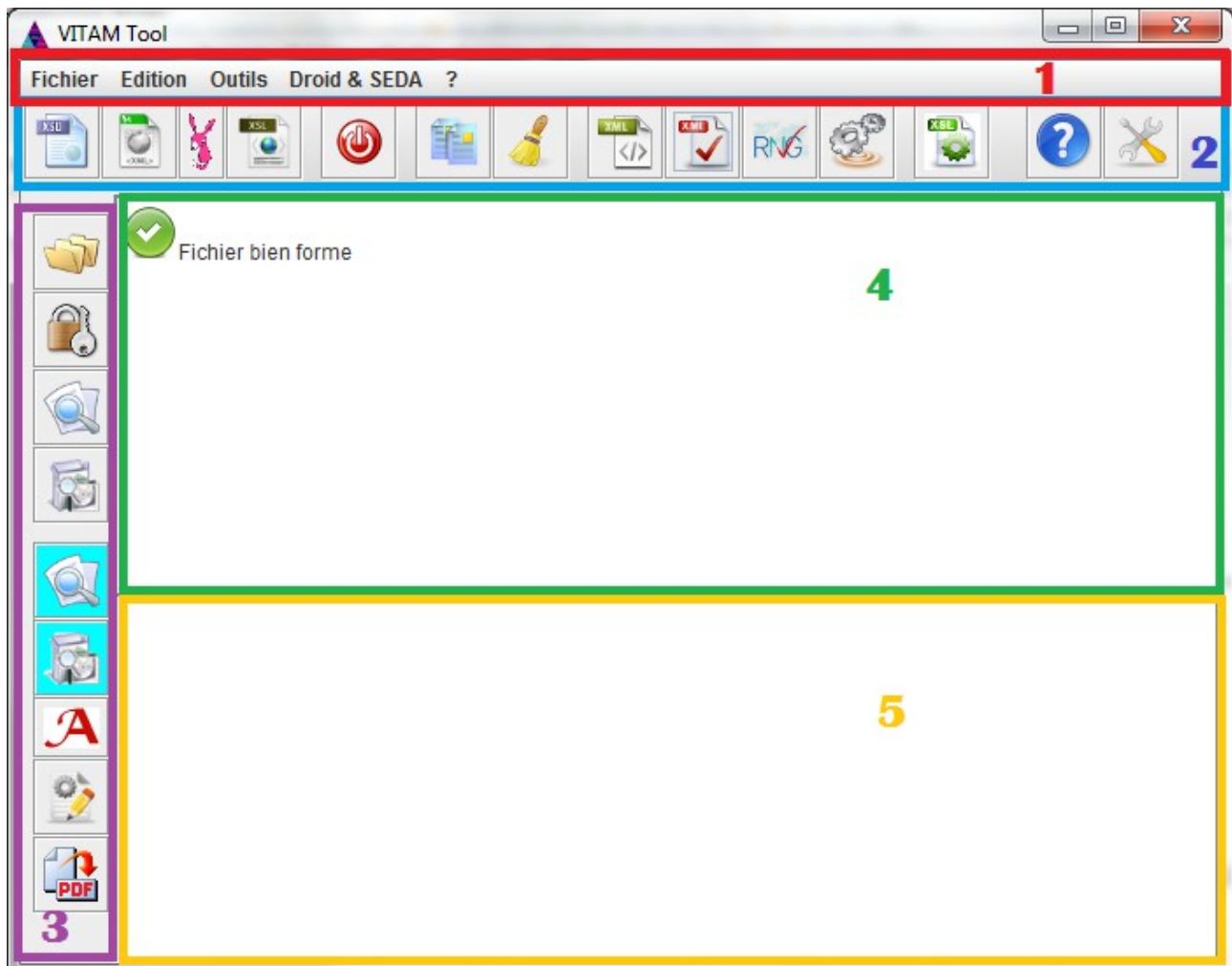
Une fois qu'un fichier XML a été chargé :



Subdivision de l'écran


L'écran est divisé comme suit :

1. En haut, le menu donnant accès à toutes les fonctions
2. En dessous, le menu boutons concernant les fonctionnalités de base sur un fichier XML
3. A droite, le menu boutons concernant les fonctionnalités avancées sur un fichier XML ou depuis un répertoire ou un document
4. La zone de texte supérieure est la zone de sortie normale (ou en mode TXT)
5. La zone de texte inférieure est la zone de sortie d'erreur
6. Un popup s'affiche dans les conditions suivantes :
 1. Demande d'accès à un fichier ou un répertoire
 2. Demande A propos
 3. Demande de la configuration (une modification enregistrera le fichier vitam-confif.xml en fonction des changements et si la case « Mise à jour du fichier de paramétrage » est coché)
 4. L'affichage d'information ou d'un résultat XML




Vitam Configuration

XML Context Tools Output

XSD 


XSD Root Namespace


Document field 

File field File attribut

MimeType attribute Format attribute

Digest field Algorithm attribut


Schematron 


XSL 


☒ Mise a jour du fichier de parametrage

Vitam Configuration

XML Context Tools Output


Droid Signature 


Droid Container 

Fits Home 

Check ☐ Archive checking ☐ Recursive checking ☒ Warn x-fmt

☒ SHA-1 Digest ☐ SHA-256 Digest ☐ SHA-512 Digest

LibreOffice/OpenOffice 

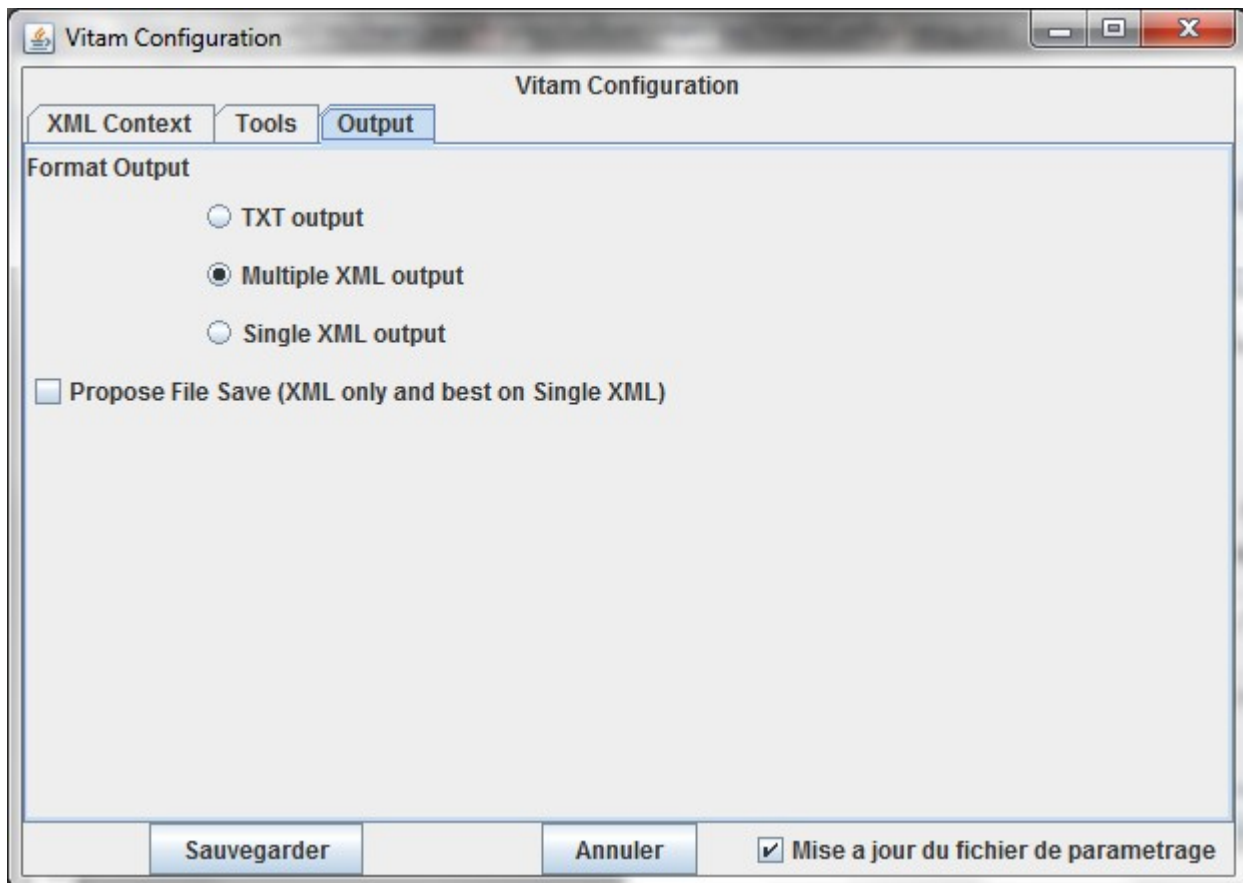
Unoconv 

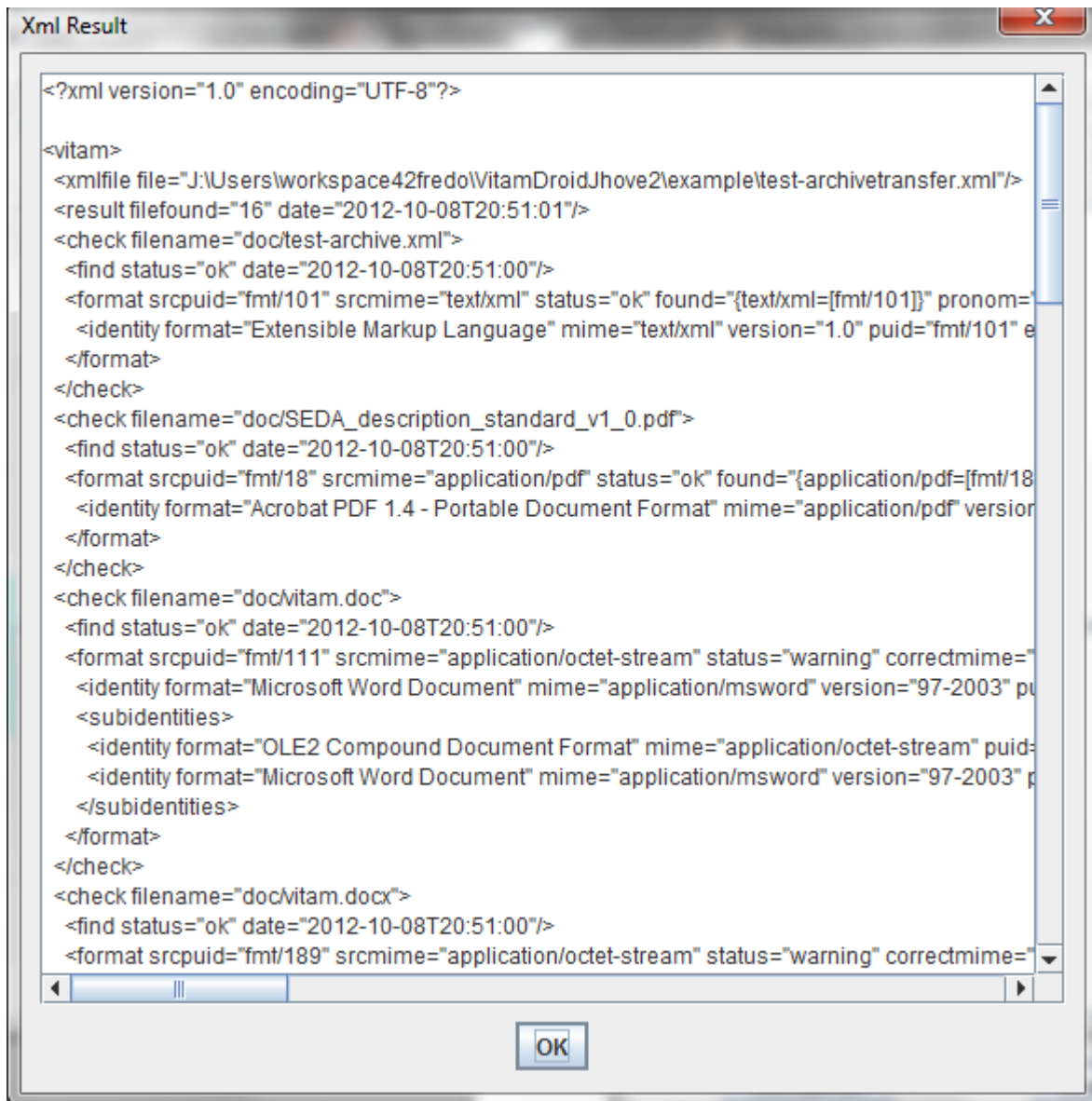
☒ Extract Keywords

to PDF/A Milliseconds per KB Keyword Limit Rank Occurence

to PDF/A Low Limit in ms Keyword Limit Word Occurence

☒ Mise a jour du fichier de parametrage





Xml Result



```
<?xml version="1.0" encoding="UTF-8"?>

<digests source="J:\Users\workspace42fredol\WitamDroidJhove2\example\doc" date="2012-10-10T20:00:00Z">
  <file filename="doc\doc.rar" sha-1="aGdrCLocvMJdf1cRt/Ga2sBDPJE=" status="ok" date="2012-10-10T20:00:00Z">
  <file filename="doc\doc.tar" sha-1="w35YNYcxe1LLf1/JRT4JjaGt3PY=" status="ok" date="2012-10-10T20:00:00Z">
  <file filename="doc\doc.tar.bz2" sha-1="RHw1jL3pO3e+2W3ZBbHRW5ZgB8=" status="ok" date="2012-10-10T20:00:00Z">
  <file filename="doc\doc.tar.gz" sha-1="xYtasvGY8N5JQEYdwZg5QEVe0zM=" status="ok" date="2012-10-10T20:00:00Z">
  <file filename="doc\doc.tgz" sha-1="xYtasvGY8N5JQEYdwZg5QEVe0zM=" status="ok" date="2012-10-10T20:00:00Z">
  <file filename="doc\doc.zip" sha-1="fynXqzi8grvB1fmJCVV6xdCH+c=" status="ok" date="2012-10-10T20:00:00Z">
  <file filename="doc\SEDA_description_standard_v1_0.pdf" sha-1="XbbauxhCwzBOIQ4wiM3i+ajUpo" status="ok" date="2012-10-10T20:00:00Z">
  <file filename="doc\test-archive.xml" sha-1="ZzOzEvlj9XwjAUAZxpgGOB5rXH4=" status="ok" date="2012-10-10T20:00:00Z">
  <file filename="doc\Witam mail.eml" sha-1="DBVGz3gYFRidZgJsJ5/j77NEk84=" status="ok" date="2012-10-10T20:00:00Z">
  <file filename="doc\vitam.doc" sha-1="FlgYpUjhVc7DC8GuBb5FGfn5Ygw=" status="ok" date="2012-10-10T20:00:00Z">
  <file filename="doc\vitam.docx" sha-1="h4Nh22jPU4AtJPghXzZr9ylL58=" status="ok" date="2012-10-10T20:00:00Z">
  <file filename="doc\vitam.gif" sha-1="SGM9cmFZVYW8puVG2khJXKCSeYM=" status="ok" date="2012-10-10T20:00:00Z">
  <file filename="doc\vitam.jpg" sha-1="U7XUW2iniLlr6wZ47SQS176V3o=" status="ok" date="2012-10-10T20:00:00Z">
  <file filename="doc\vitam.odp" sha-1="F53WIW5xW08T7eKfLuYXQPfEi18=" status="ok" date="2012-10-10T20:00:00Z">
  <file filename="doc\vitam.ods" sha-1="TZVlpojmgOF7WQMPx0+ylsrxUU4=" status="ok" date="2012-10-10T20:00:00Z">
  <file filename="doc\vitam.odt" sha-1="3SJh6xBnf1Pw7ZcBwrmDJnyEVKw=" status="ok" date="2012-10-10T20:00:00Z">
  <file filename="doc\vitam.png" sha-1="TaUF6QuihT0Y3bJxzyefy67O0uY=" status="ok" date="2012-10-10T20:00:00Z">
  <file filename="doc\vitam.ppt" sha-1="4gNANlfrg+2n3gcNI1u9RgAtle4=" status="ok" date="2012-10-10T20:00:00Z">
  <file filename="doc\vitam.pptx" sha-1="3L5a8WjJviwAXp2ZIGXlx+by9FA=" status="ok" date="2012-10-10T20:00:00Z">
  <file filename="doc\vitam.tif" sha-1="NS/5cD6+TSsB5BONkNOYm6FfWF4=" status="ok" date="2012-10-10T20:00:00Z">
  <file filename="doc\vitam.xls" sha-1="ZaM5QCkZomg+ZM3OWsbgcGDIS2Q=" status="ok" date="2012-10-10T20:00:00Z">
  <file filename="doc\vitam.xlsx" sha-1="GGqgHx/VL7FI8noAZaVhK9CumFg=" status="ok" date="2012-10-10T20:00:00Z">
  <file filename="doc\vitamtool.pdf" sha-1="YXegJM7WJlZgW2U3g4JuU1bXZOE=" status="ok" date="2012-10-10T20:00:00Z">
</digests>
```

OK