E-Paye

Extracteur des informations de la base SILAE

Table des matières

[Objectif 3](#_Toc535566926)

[Prerequis 3](#_Toc535566927)

[Outils logiciels pour utiliser l’extracteur 3](#_Toc535566928)

[Données SILAE 3](#_Toc535566929)

[Principe d’utilisation 3](#_Toc535566930)

[Fichier de configuration 3](#_Toc535566931)

[Emplacement et nom du fichier 3](#_Toc535566932)

[Structure du fichier de configuration 3](#_Toc535566933)

[Contenu du fichier de configuration (exemple) : 3](#_Toc535566934)

[Description des clés utilisées : 4](#_Toc535566935)

[Déploiement 4](#_Toc535566936)

[Extraction 4](#_Toc535566937)

[Fichiers extraits 5](#_Toc535566938)

[Principe de fonctionnement 5](#_Toc535566939)

[Schema fonctionnel de l’extracteur 5](#_Toc535566940)

[ContextE défini dans le projet 5](#_Toc535566941)

[Jobs utilisés : séquence 5](#_Toc535566942)

[Job de configuration « general\_context\_loader » 6](#_Toc535566943)

[Job d’orchestration 6](#_Toc535566944)

[Jobs d’extraction des données 7](#_Toc535566945)

[Construction de l’extracteur 8](#_Toc535566946)

[AnnexeS 8](#_Toc535566947)

[Exemple de fichier de configuration 8](#_Toc535566948)

[Script SQl de création de la base de données 9](#_Toc535566949)

[Références 9](#_Toc535566950)

[Schéma relationel partiel de la base SILAE 9](#_Toc535566951)

# Objectif

L’objectif de cet outil est de pouvoir extraire les informations des clients à partir de la base de données SILAE afin d’alimenter l’outils d’analyse de données YOUREPORT.

Afin de réaliser cette opération, nous utilisons un outil spécifique, TALEND, de génération de fichiers CSV à partir de la base de données SILAE.

Cet outil permet de façonner les extractions de manière à être compatible avec les modèles existants déjà utilisés sur le logiciel e-Paye V1.

# Prerequis

## Outils logiciels pour utiliser l’extracteur

Les différents outils logiciels nécessaires sur l’hôte sont les suivants :

* Un moteur de base de données MySQL version 5.7.18 au minimum.
* Le « framework » JAVA JRE en version 1.8 au minimum
* Pour le développement et les modifications, framework TALEND en version 1.7 au minimum.

## Données SILAE

La base de données SILAE du client au format SQL.

# Principe d’utilisation

## Fichier de configuration

Le fichier de configuration permet de paramétrer l’extracteur.

### Emplacement et nom du fichier

Sa localisation sur l’hôte doit être située **OBLIGATOIREMENT** en :

* Sur UNIX : « /talend/contextConfiguration »
* Sur Windows : « c:\talend\contextConfiguration »

Le fichier doit **obligatoirement** porter le nom : « Default.General.cfg »

### Structure du fichier de configuration

La forme générale du fichier est structurée par le couple : clé=valeur.

Chaque couple est situé sur une ligne différente du fichier.

### Contenu du fichier de configuration (exemple) :

BaseDirectoryForSql=C:/Users/lauthier/TOS\_DI-Win32-20181026\_1147-V7.1.1/test\_export\_job

ImportSqlScript=create\_database\_silae\_auto\_import.sql

DataBaseHost=localhost

DataBasePort=3306

DataBaseName=silae

DataBaseUser=root

DataBasePasswordUser=laurent

ExtractDirectory=/extract\_silae

ArchiveName=

PathPrefixInZip=export/home/dev\_epaye/myreportbdd

TmpDirectory=/tmp\_talend

NumeroDossier=6030909,6028381,6033158

### Description des clés utilisées :

* BaseDirectoryForSql : permet de localiser sur l’hôte le fichier SQL représentant la base SILAE du client ainsi que le script de création de la base.
* ImportSqlScript : défini le nom du script SQL permettant de créer la base SILAE à partir du script d’origine SQL SILAE.
* DataBaseHost : défini l’hôte qui supporte MySql
* dataBasePort : défini le port MySql utilisé
* DataBasename : défini le nom de la base créée au sein de MySql
* DataBaseUser : défini le nom de l’utilisateur MySql
* DataBasePasswordUser : défini le mot de passe MySql.
* ExtractDirectory : défini l’emplacement sur l’hôte où seront stockés les archives d’extraction
* ArchiveName : défini un préfixe générique commun à toutes les extractions.
* PathPrefixInZip : défini un chemin utilisé au sein de l’archive pour atteindre les fichiers CSV.
* TmpDirectory : défini un répertoire de stockage temporaire utilisé lors des phases d’extractions.
* NumeroDossier : défini la liste des numéros de dossiers à extraire, le séparateur étant la virgule.

## Déploiement

1. Décompresser le fichier « extractSilaeToYouReport.zip» dans un répertoire adéquate de l’hôte.
2. Créer le fichier de configuration
3. Créer les répertoires définis dans le fichier de configuration
4. Copiez le fichier de configuration « Default.General.cfg » de la racine de l’archive extraite vers « /talend/contextConfiguration/ »
5. Copier le fichier « create\_database\_silae\_auto\_import.sql » de la racine de l’archive vers l’emplacement défini par la clé « BaseDirectoryForSql ».
6. Copiez le fichier client de la base de données SILAE qui doit s’intituler **obligatoirement** « silae\_dom.sql » vers l’emplacement défini par la clé « BaseDirectoryForSql ».
7. Définissez les numéros de dossiers à extraire en renseignant les valeurs de la clé « NumeroDossier ».

## Extraction

Afin d’extraire les fichiers à partir de l’outils d’extraction il faut lancer le script suivant :

* Sous UNIX : extractSilaeToYouReport.sh
* Sous Windows : extractSilaeToYouReport.bat

Les archives extraites sont disponibles à l’emplacement de la clé «ExtractDirectory» .

## Fichiers extraits

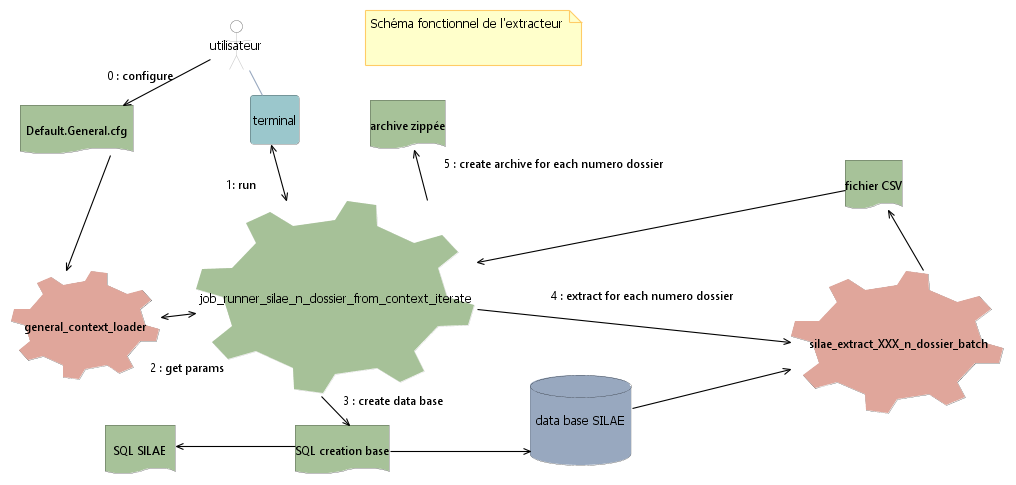
Le format du nom des archives extraites est de la forme : <numeroDossier>\_AAAAMM

Chaque archive contient l’ensemble des fichiers CSV avec éventuellement un préfixe de chemin défini par la clé PathPrefixInZip .

# Principe de fonctionnement

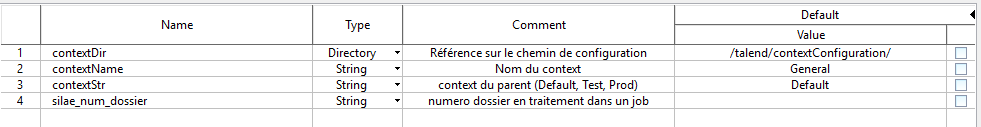
Afin de comprendre correctement l’ensemble des explications ci-dessous, il est nécessaire de connaitre la philosophie et le principe de fonctionnement de l’outil « TALEND ».

## Schema fonctionnel de l’extracteur



## ContextE défini dans le projet

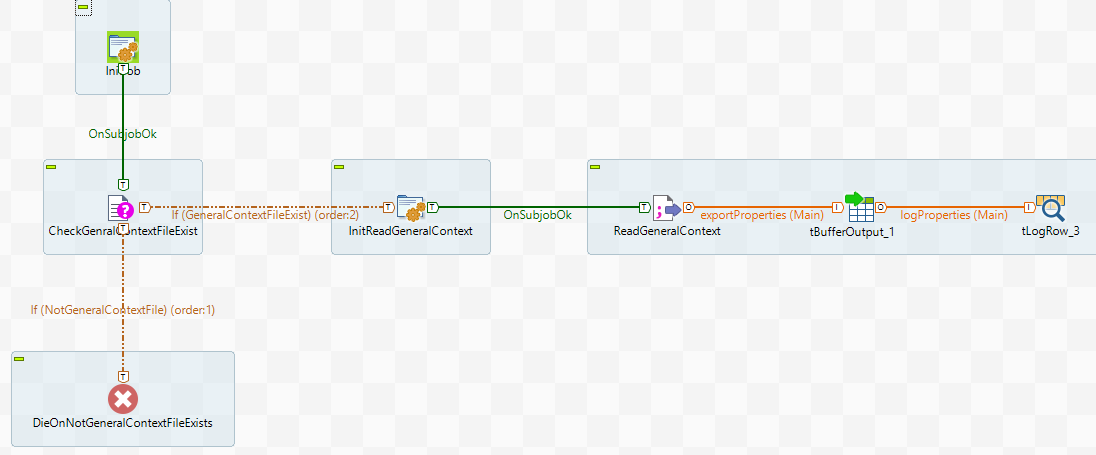
Le contexte général défini est le suivant :



## Jobs utilisés : séquence

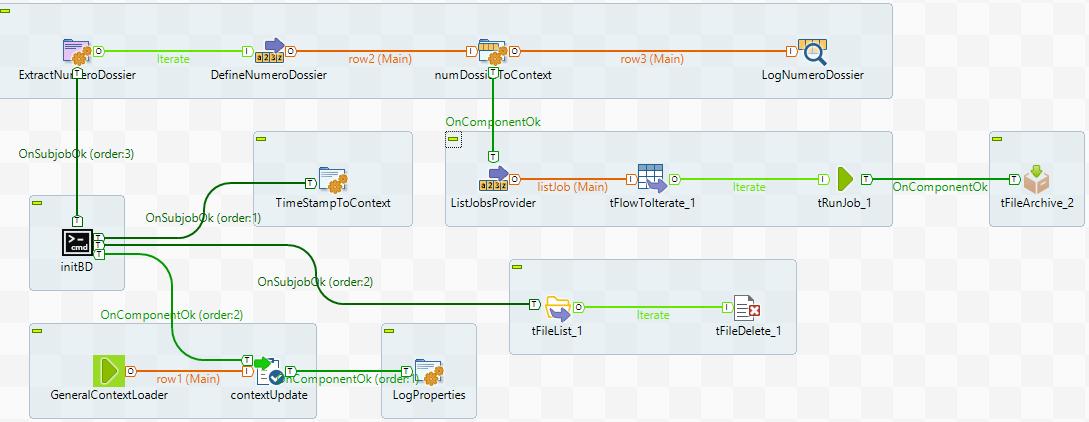
1. general\_context\_loader: permet de charger le contexte avec les valeurs contenues dans le fichier de paramétrage.
2. job\_runner\_silae\_n\_dossier\_from\_context\_iterate : permet
   1. De configurer le contexte à partir du job « general\_context\_loader »
   2. D’initialiser la base de données ainsi que nettoyer le répertoire temporaire
   3. De lancer l’ensemble des jobs d’extraction
   4. De créer les fichiers archives à partir des fichiers CSV.
3. silae\_extract\_XXXXXXX\_n\_dossier\_batch : permettent d’extraire les information de la base de données SILAE en créant un fichier CSV spécifique

## Job de configuration « general\_context\_loader »



Dans ce job d’initialisation si le fichier de paramétrage est introuvable le processus d’extraction s’interrompt.

## Job d’orchestration



Ce job obéit à l’algorithme suivant :

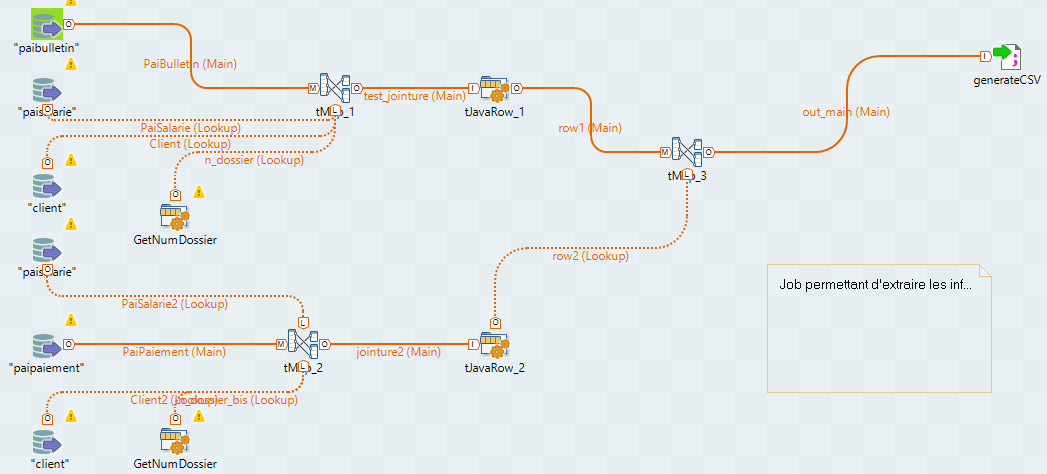
1. Lance la récupération des paramètres pour initialiser le contexte.
2. Supprime les fichiers et répertoires dans le répertoire temporaire défini
3. Fabrique le « timeStamp » pour les archives
4. Pour chaque numéro de dossier
   1. Lance les jobs d’extraction définis dans le ListJobsProvider
   2. A chaque fin de job enrichi l’archive avec le fichier CSV

Quelques points de détails :

* Le composant InitBD joue le script de création de la base de données. C’est lui qui déclenche le reste du processus d’extraction.
* Le composant ExtractNumeroDossier permet d’extraire les uns après les autres les numéros de dossiers. C’est lui qui est à l’origine des itérations.

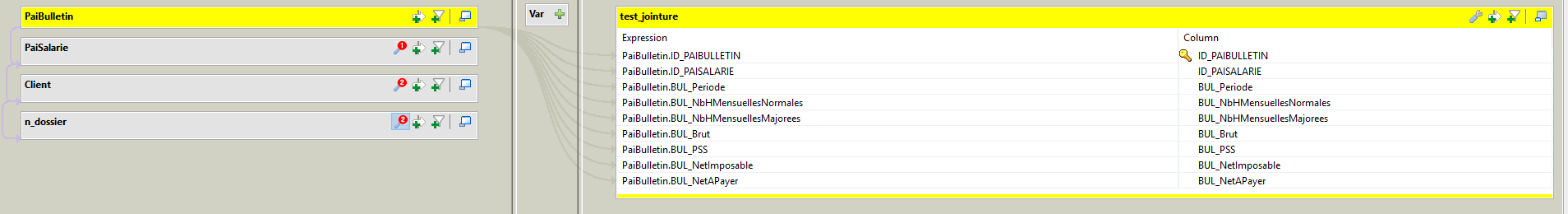
## Jobs d’extraction des données

Voici l’exemple du job d’extraction des informations de bulletins suivant un numéro de dossier :

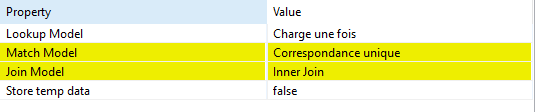


Dans ce job on rencontre de gauche à droite :

* Les tables de la base de données impliquées dans cette extraction.
* Le composant GetNumDossier responsable de la mise à disposition du numéro de dossier à traiter.
* Les composants « tMap » qui permettent de faire des jointures entre les tables. Deux types de liens sont utilisés :
  + Le lien « Main » qui permet de relier la table principale de jointure.
  + Les liens « LookUp » qui permettent de relier les tables de recherche de concordance. A noter que l’on peut évidemment chaîner les jointures sur les liens « LookUp »
  + Un exemple de configuration de tMap1 :



Les jointures sont configurées de la manière suivante :



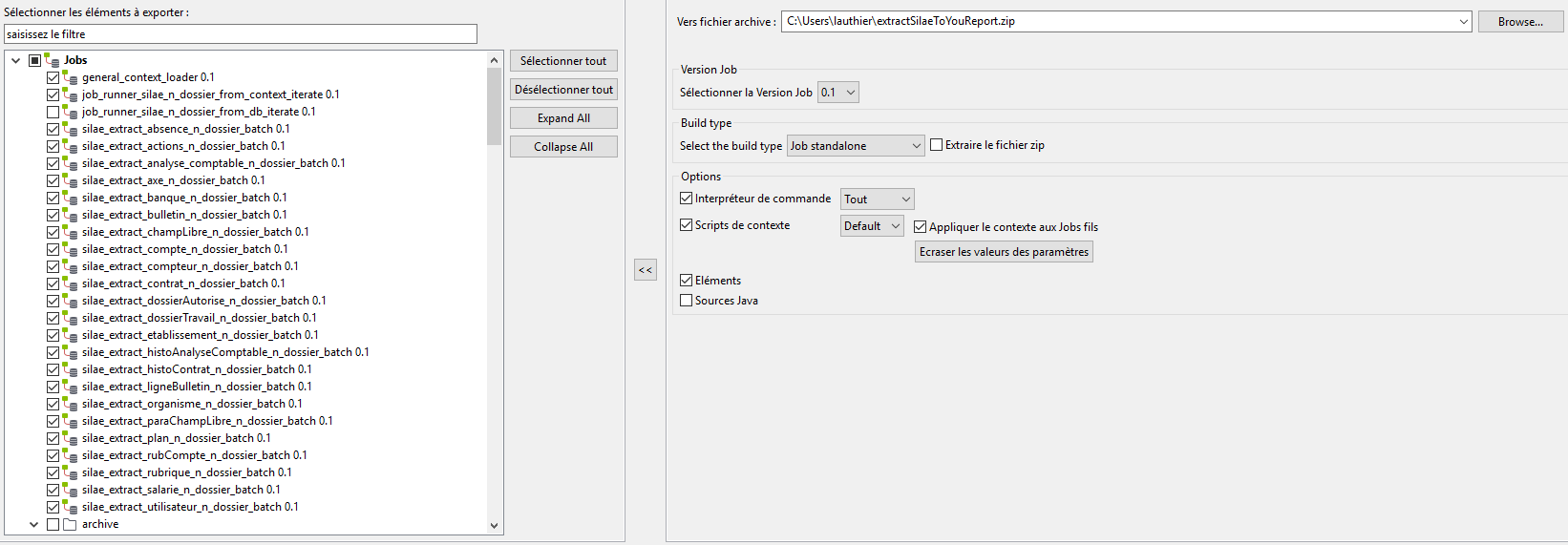
* Le dernier composant « tMap » est précédé de composants « tJavaRow » qui permettent de définir le flux conservé à l’issue des jointures de « tMap » ; en effet, la connexion directe entre les « tMap » ne permet pas de fournir le résultat final.
* Le composant « GenerateCSV » qui fabrique le fichier CSV avec le bon nom.

L’ensemble des autres jobs d’extraction sont similaires dans leur construction.

# Construction de l’extracteur

Afin de construire l’outil à partir du framework « TALEND », procédez comme suit :

1. Ouvrir le projet.
2. Sélectionner le job job\_runner\_silae\_n\_dossier\_from\_context\_iterate
3. Cliquez droit et sélectionnez « construire le job »
4. Complétez le formulaire en vous inspirant de la copie d’écran ci-dessous (adaptez le chemin qui supporte la génération du zip)



1. Cliquez Finish

# AnnexeS

## Exemple de fichier de configuration

Nom du fichier Default.General.cfg :

BaseDirectoryForSql=C:/Users/lauthier/TOS\_DI-Win32-20181026\_1147-V7.1.1/test\_export\_job

ImportSqlScript=create\_database\_silae\_auto\_import.sql

DataBaseHost=localhost

DataBasePort=3306

DataBaseName=silae

DataBaseUser=root

DataBasePasswordUser=laurent

ExtractDirectory=/extract\_silae

ArchiveName=

PathPrefixInZip=export/home/dev\_epaye/myreportbdd

TmpDirectory=/tmp\_talend

NumeroDossier=6030909,6028381,6033158

## Script SQl de création de la base de données

nom du script « create\_database\_silae\_auto\_import.sql »

CREATE DATABASE IF NOT EXISTS silae\_auto\_import CHARACTER SET = 'utf8' COLLATE = 'utf8\_general\_ci';

use silae\_auto\_import;

source ./silae\_dom.sql;

# Références

* <https://www.talend.com/products/data-integration/data-integration-open-studio/?utm_medium=talendcommunity&utm_source=community_navmenu>
* https://community.talend.com/
* <https://help.talend.com/>
* <https://www.talendbyexample.com/>

# Schéma relationel partiel de la base SILAE

