

# **INTÉGRATION DE DONNÉES ETL TALEND**

**AUTEUR : JONATHAN FONTANEL**

**DATE : 11/12/2023**

**TP ISIMA ZZ2 F4 BASES DE DONNÉES ET FOUILLES DE DONNÉES**



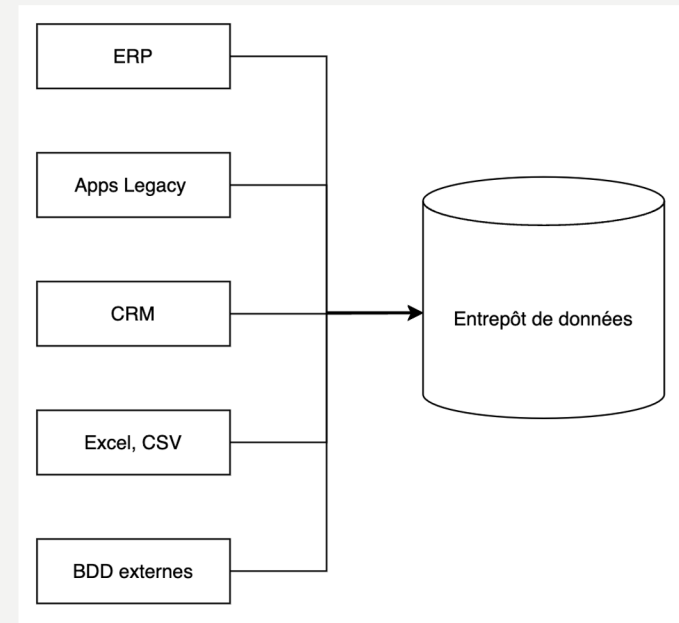
# **ETL ET PRÉSENTATION DES OUTILS**

# ETL : EXTRACT, TRANSFORM LOAD

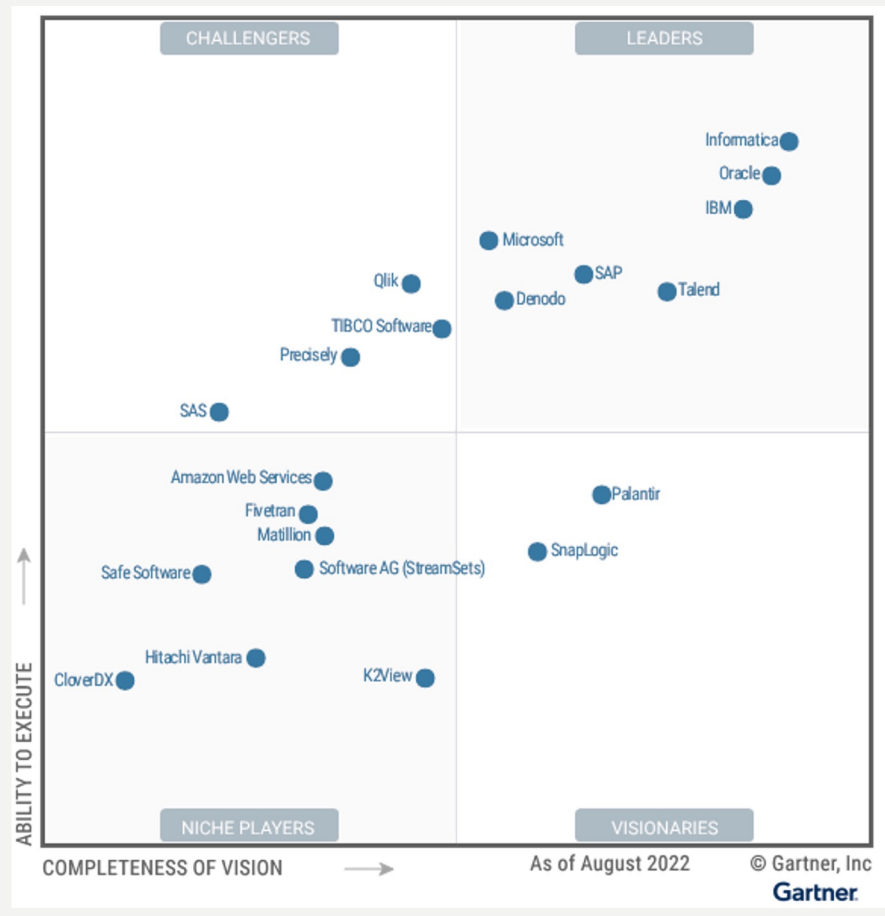
- Extraction de données de tous types (Excel, fichier plat ou CSV, BDD, etc...)
- Transformation de données (Création de colonnes, changement de type, Jointure de tables)
- Chargement de données (Chargement dans l'entrepôt de données ou des fichiers)

# PRÉSENTATION DE TALEND OPEN STUDIO

- But d'un ETL : l'intégration de données
  - Intégration opérationnelle :
    - Partage de données entre applications avec transformation ou non
    - Flux temps réels entre applications ;
    - Réplication de base de données ;
    - Migration de données ;
    - Lien avec des partenaires (fournisseurs clients, ...)
    - ...
  - Intégration décisionnelle :
    - Charger des schémas du datawarehouse
    - Charger des datalakes
    - Transformation des données entre les différentes couches du datawarehouse
    - ...



# MAGIC QUADRANT FOR DATA INTEGRATION TOOLS



# MAGIC QUADRANT FOR INTEGRATION PLATFORM AS A SERVICE



# PRÉSENTATION DE TALEND OPEN STUDIO EN 2022 ET 2023

- Talend Open Studio for Data Integration : Outil Open Source pour l'intégration de données

- Versions payantes :

## Stitch

Gain de temps pour les analystes grâce à l'ingestion de données sans code

Contactez le service commercial

Démo à la demande

- Exploitez des insights provenant des 140 principales sources de données du secteur
- Exploitez les data warehouses cloud leaders du marché
- Structurez la donnée pour les principaux outils d'analyse
- Orchestrez votre pipeline de données en fonction de vos besoins
- Tirez profit d'une connectivité avancée et réduisez vos coûts de transfert
- Bénéficiez du support adapté à votre activité

## Data Management Platform

Le point de départ idéal pour les professionnels des données et leurs équipes

Contactez le service commercial

Démander une démo

- ✓ Flexibilité d'intégration multi-cloud, cloud hybride et on-premise
- ✓ Travail collaboratif
- ✓ Profilage des données pour l'identification des problèmes de qualité
- ✓ Recettes de transformation partageables et réutilisables pour des solutions rapides aux problèmes de qualité les plus courants
- ✓ Fonctionnalités de préparation des données en libre-service

Connecteurs spécialisés pour déplacer de gros volumes de données

Fonctionnalités et capacités supplémentaires pour les données non structurées

Traitement batch avec Spark pour un mouvement parfait des données

Capacité à concevoir, tester et créer des API pour le partage de données

Microservices flexibles pour des cycles de développement optimaux

Intégration application-à-application

## Big Data Platform

Analytiques avancées pour des initiatives entre équipes

Contactez le service commercial

Démander une démo

- ✓ Flexibilité d'intégration multi-cloud, cloud hybride et on-premise
- ✓ Travail collaboratif
- ✓ Profilage des données pour l'identification des problèmes de qualité
- ✓ Recettes de transformation partageables et réutilisables pour des solutions rapides aux problèmes de qualité les plus courants
- ✓ Fonctionnalités de préparation des données en libre-service

Connecteurs spécialisés pour déplacer de gros volumes de données

Fonctionnalités et capacités supplémentaires pour les données non structurées

Traitement batch avec Spark pour un mouvement parfait des données

Capacité à concevoir, tester et créer des API pour le partage de données

Microservices flexibles pour des cycles de développement optimaux

Intégration application-à-application

## Data Fabric

La solution d'intégration et de partage de données la plus complète pour toute l'entreprise

Contactez le service commercial

Démander une démo

- ✓ Flexibilité d'intégration multi-cloud, cloud hybride et on-premise
- ✓ Travail collaboratif
- ✓ Profilage des données pour l'identification des problèmes de qualité
- ✓ Recettes de transformation partageables et réutilisables pour des solutions rapides aux problèmes de qualité les plus courants
- ✓ Fonctionnalités de préparation des données en libre-service

Connecteurs spécialisés pour déplacer de gros volumes de données

Fonctionnalités et capacités supplémentaires pour les données non structurées

Traitement batch avec Spark pour un mouvement parfait des données

Capacité à concevoir, tester et créer des API pour le partage de données

Microservices flexibles pour des cycles de développement optimaux

Intégration application-à-application



**PLACE À LA  
PRATIQUE**





**TP1**

**TALEND OPEN STUDIO**

# TO DO

- 1<sup>er</sup> job pour lire une table et ajouter un champ
  - Création d'un projet
  - Création d'une connexion à une base MySQL
  - Création d'un premier job
  - Ajout du composant tMap
  - Ajout d'un champ calculé à partir d'un champ présent
  - Exécution du job avec un champ supplémentaire

# PRÉREQUIS

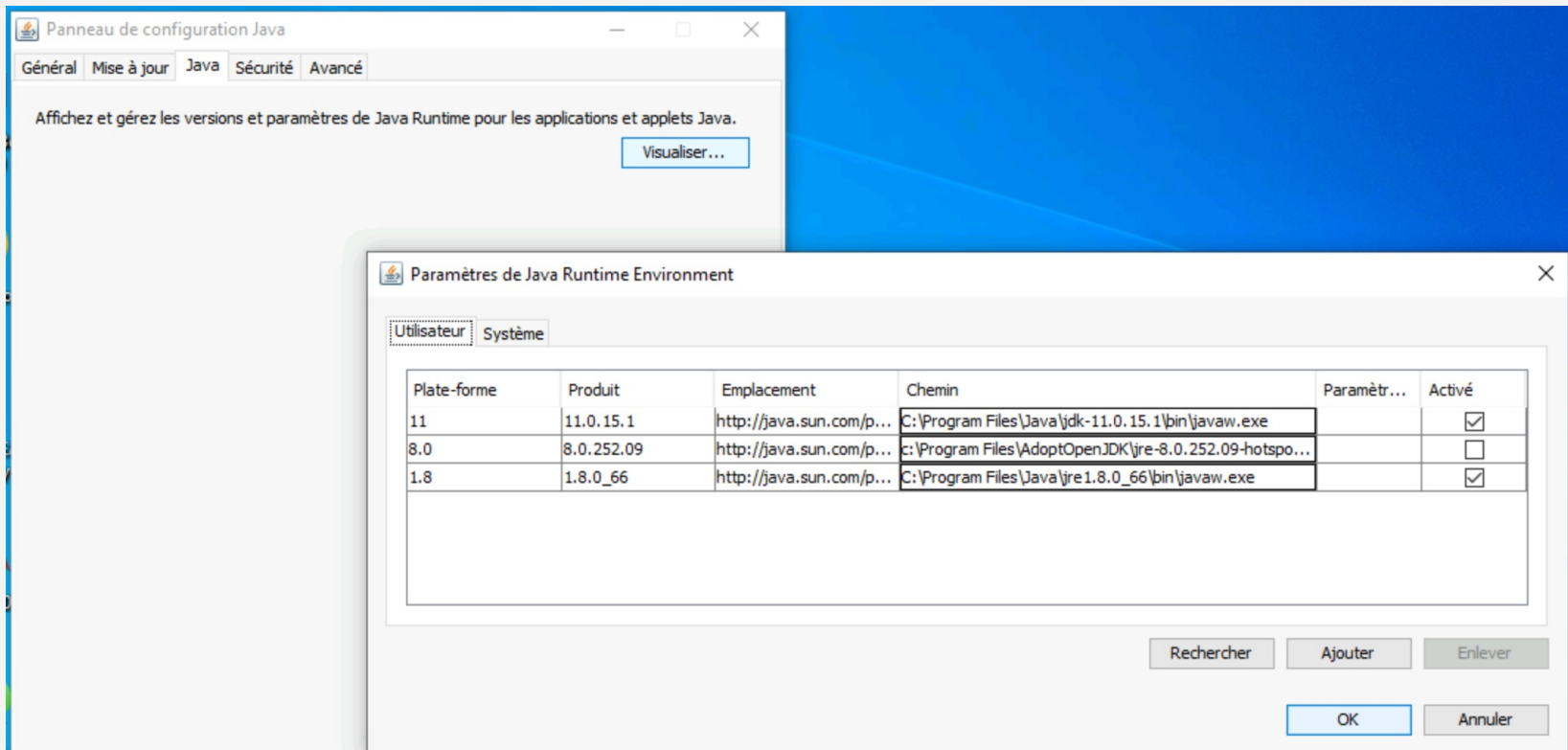
- ~~Installer MySQL Workbench + MySQL Community Server~~
  - ~~Lien de téléchargement : <https://dev.mysql.com/downloads/>~~
  - ~~Paramétrage en mode « Use Legacy Password Encryption » pour les mots de passe lors de l'installation de MySQL Community Server~~
- Installation de la base « gestion\_livres »
  - **Lien de la base Gestion\_livres**
  - Exécuter les deux scripts SQL dans le schéma « gestion\_livres » que vous aurez créé préalablement
- Récupération des paramètres de connexion à la base (IP, PORT, nom de la base de données, mot de passe d'accès pour le compte ROOT)
- Vérification avec une requête
- ~~Installation de Talend à partir du fichier suivant : **Lien Talend**~~

# PRÉREQUIS

- Vérification de MySQL Workbench
- Installation de la base « gestion\_livres »
- Récupération des paramètres de connexion à la base
- Vérification avec une requête

# ATELIERS SUR TALEND OPEN STUDIO

- Vérifier le JDK sur la machine et si nécessaire installer le JDK 11 :



# ATELIERS SUR TALEND OPEN STUDIO

- Configuration du JDK :

The screenshot shows the Windows System Properties dialog box with the 'Advanced' tab selected. The 'Environment Variables' button is highlighted. Two sub-dialogs are open:

- Variables d'environnement**: Shows user variables for 'jofontan' and system variables.
- Modifier la variable d'environnement**: Shows the list of system variables with 'C:\Program Files\Java\jdk-11.0.15.1\bin' selected.

**Variables utilisateur pour jofontan**

Variable	Valeur
OneDrive	C:\Users\jofontan\OneDrive
Path	C:\Users\jofontan\AppData\Local\Microsoft\WindowsApps;
TEMP	C:\Users\jofontan\AppData\Local\Temp
TMP	C:\Users\jofontan\AppData\Local\Temp

**Variables système**

Variable	Valeur
MAVEN_HOME	C:\Program Files\apache-maven-3.6.3
NUMBER_OF_PROCESSORS	4
OS	Windows_NT
Path	C:\Program Files\Java\jdk-11.0.15.1\bin;C:\Program Files\Common...
PATHEXT	.COM;.EXE;.BAT;.CMD;.VBS;.VBE;.JS;.JSE;.WSF;.WSH;.MSC
PROCESSOR_ARCHITECTURE	AMD64
PROCESSOR_IDENTIFIER	Intel64 Famili 6 Model 158 Steppin 9. GenuineIntel

**Modifier la variable d'environnement**

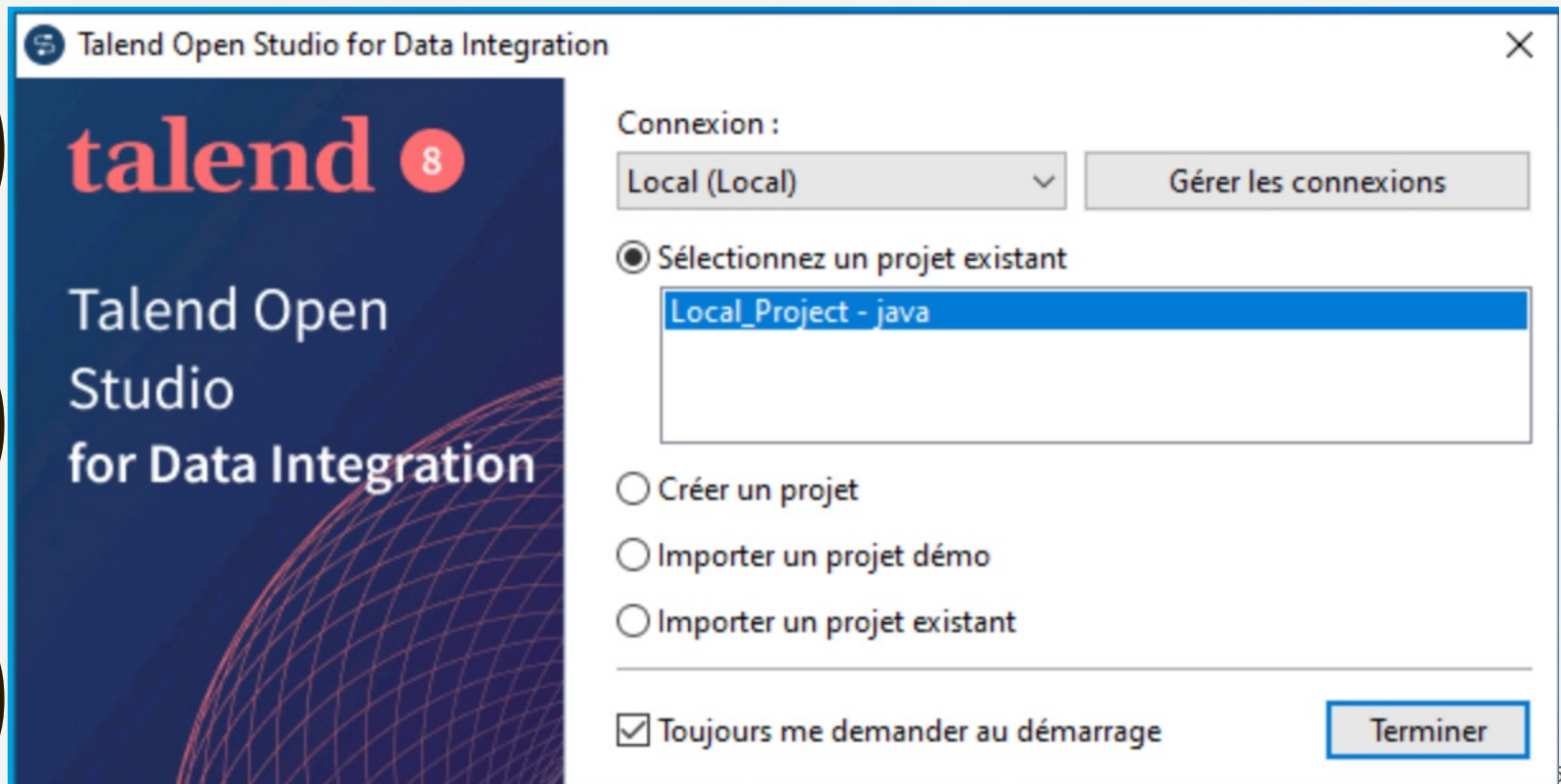
Selected variable: C:\Program Files\Java\jdk-11.0.15.1\bin

Other variables in the list:

- C:\Program Files\Common Files\Oracle\Java\javapath
- c:\Program Files\AdoptOpenJDK\jre-8.0.252.09-hotspot\bin
- c:\Program Files (x86)\AdoptOpenJDK\jre-8.0.252.09-hotspot\bin
- C:\Program Data\Oracle\Java\javapath
- %SystemRoot%\system32
- %SystemRoot%
- %SystemRoot%\System32\Wbem
- %SYSTEMROOT%\System32\WindowsPowerShell\v1.0\
- %SYSTEMROOT%\System32\OpenSSH\
- C:\Program Files (x86)\wapt
- C:\Program Files\dotnet\
- C:\Program Files\Microsoft SQL Server\130\Tools\Binn\
- C:\Program Files\Microsoft SQL Server\Client SDK\ODBC\170\T...
- C:\Program Files\Microsoft VS Code\bin
- C:\Program Files (x86)\Microsoft SQL Server\110\DTS\Binn\
- C:\Program Files (x86)\Microsoft SQL Server\120\DTS\Binn\
- C:\Program Files (x86)\Microsoft SQL Server\130\DTS\Binn\
- C:\Program Files (x86)\Microsoft SQL Server\140\DTS\Binn\
- C:\Program Files (x86)\Microsoft SQL Server\150\DTS\Binn\

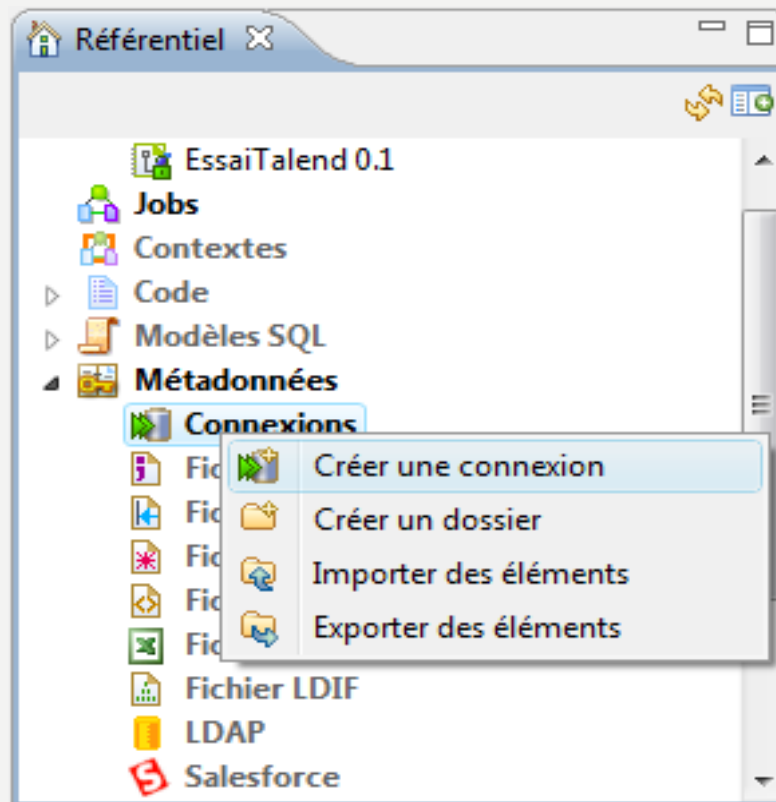
# ATELIERS SUR TALEND OPEN STUDIO

- Création d'un projet :



# ATELIERS SUR TALEND OPEN STUDIO

- Création d'une connexion à une base MySQL :





# ATELIERS SUR TALEND OPEN STUDIO

- Création d'une connexion à une base MySQL :

Connexion à la base de données

**Mettre à jour la connexion à une base de données - Etape 2/2**

*i* Vous devez cliquer sur le bouton Vérifier afin de vérifier les paramètres de la base de données.

Type de Bdd

Version de la base de données

Chaîne de caractères de connexion

Identifiant

Mot de passe

Serveur

Port

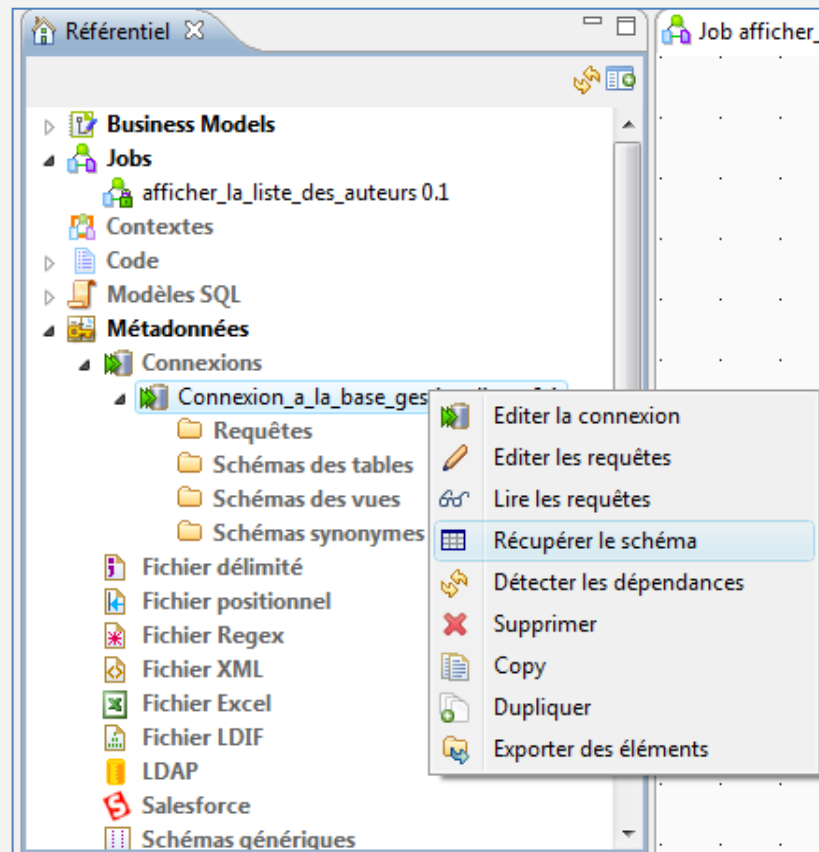
Base de données

Paramètres supplémentaires

Tester la connexion

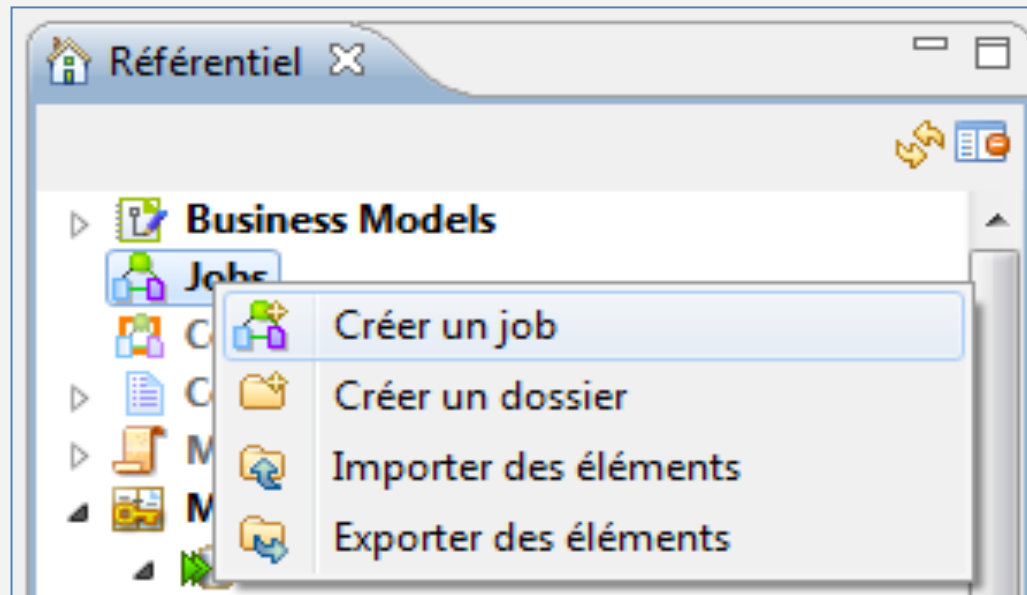
# ATELIERS SUR TALEND OPEN STUDIO

- Création d'une connexion à une base MySQL :
  - Récupération du schéma des tables



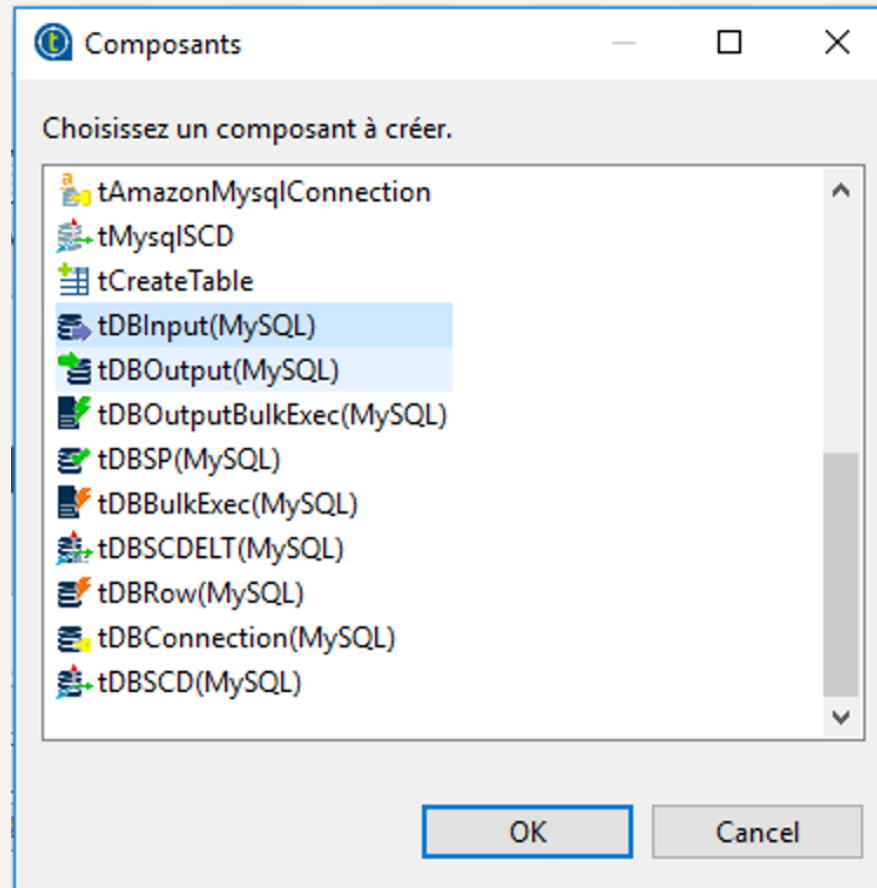
# ATELIERS SUR TALEND OPEN STUDIO

- Création d'un premier job :



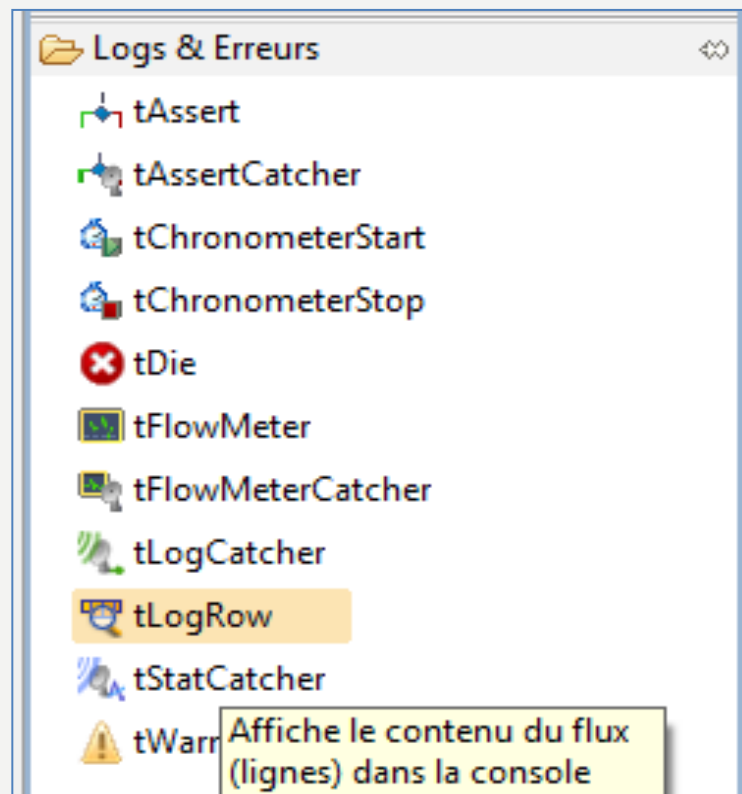
# ATELIERS SUR TALEND OPEN STUDIO

- Création d'un premier job : ajout de la connexion
  - Faire glisser le tDBInput (MySQL) dans la partie Design



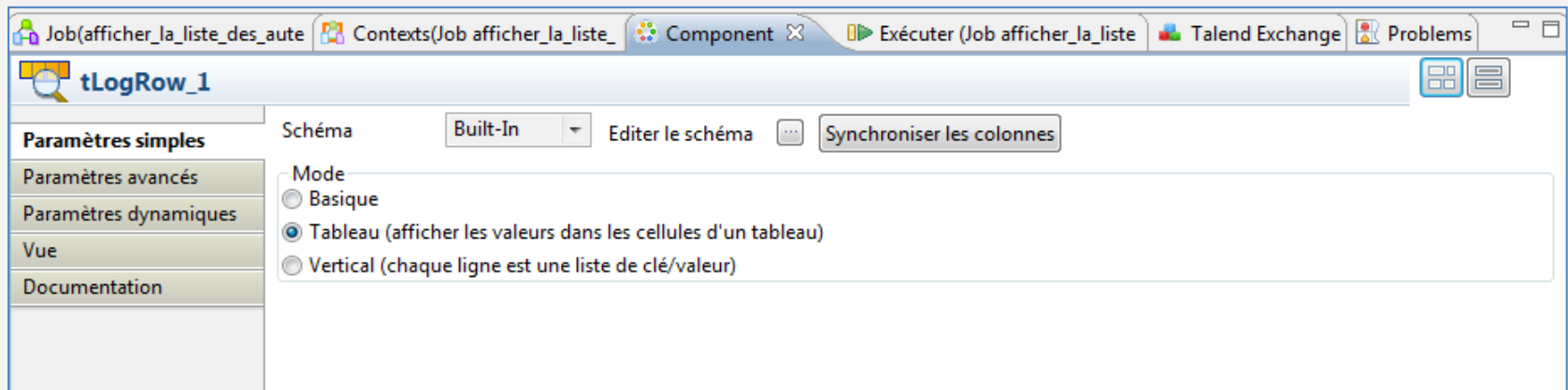
# ATELIERS SUR TALEND OPEN STUDIO

- Création d'un premier job : ajout d'un composant de log



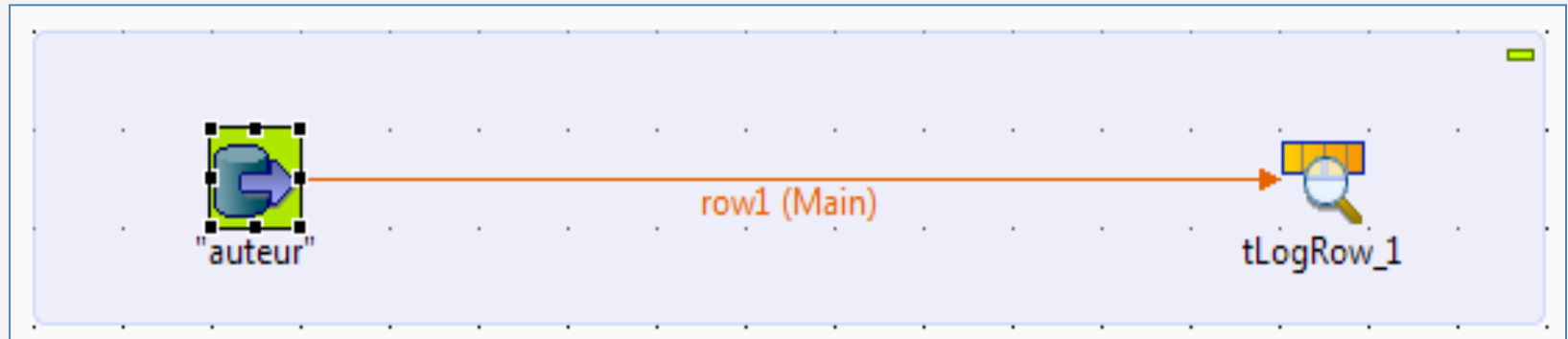
# ATELIERS SUR TALEND OPEN STUDIO

- Création d'un premier job : configuration du composant de log



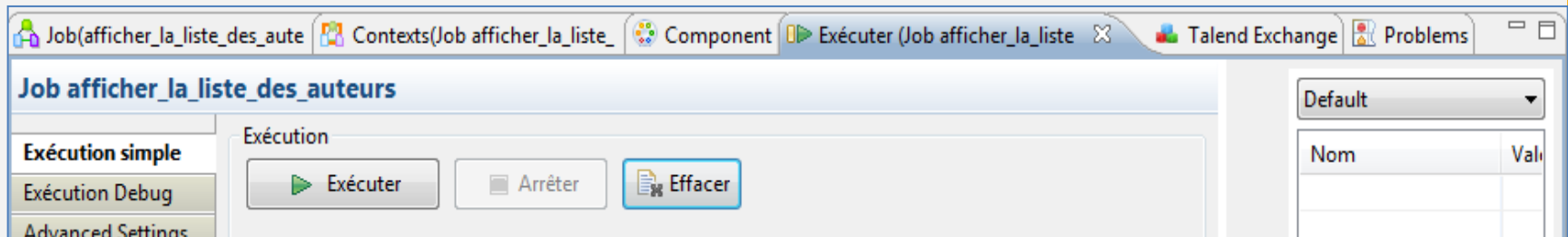
# ATELIERS SUR TALEND OPEN STUDIO

- Création d'un premier job : liaison des composants



# ATELIERS SUR TALEND OPEN STUDIO

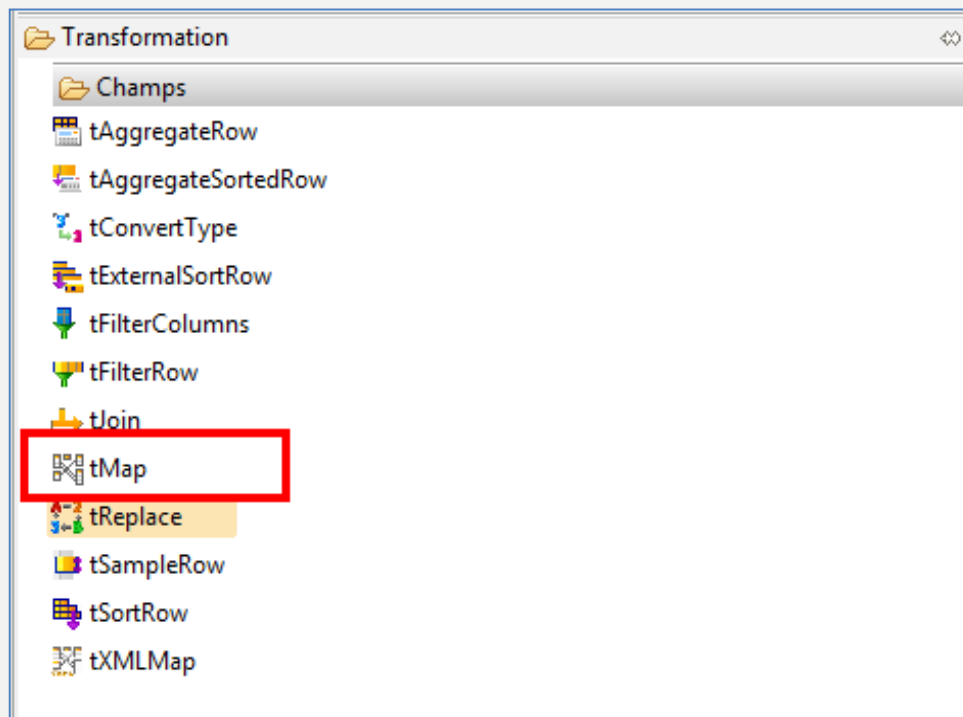
- Création d'un premier job : exécution du job





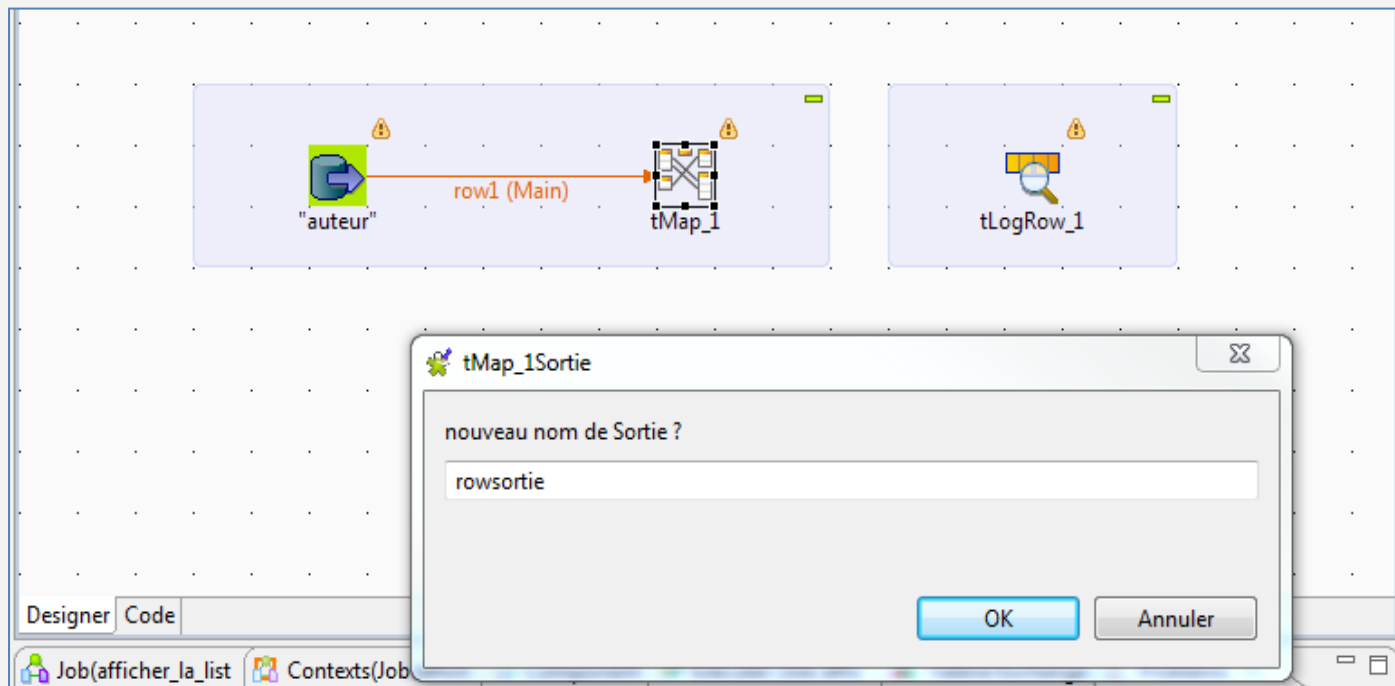
# ATELIERS SUR TALEND OPEN STUDIO

- Modification de ce job : ajout d'un composant d'aiguillage → tMap



# ATELIERS SUR TALEND OPEN STUDIO

- Modification de ce job : liaison avec les autres composants



# ATELIERS SUR TALEND OPEN STUDIO

- Modification de ce job :
  - Ouvrir le tMap
  - Puis appuyer sur mapping auto en haut à droite
  - Exécuter le job à nouveau (normalement il n'y a pas de changement)

The screenshot shows the Talend Open Studio interface with the 'Mapping auto' window open. The window displays two tables: 'row1' on the left and 'out' on the right. The 'row1' table has columns: NUMERO\_A, NOM, PRENOM, DOMICILE. The 'out' table has columns: NUMERO\_A, NOM, PRENOM, DOMICILE. The 'out' table also has an 'Expression' column with values: row1.NUMERO\_A, row1.NOM, row1.PRENOM, row1.DOMICILE. The 'Mapping auto' window is open, showing the mapping between the two tables. The 'row1' table is selected, and the 'out' table is also selected. The 'Mapping auto' window is open, showing the mapping between the two tables. The 'row1' table is selected, and the 'out' table is also selected.

Colonne	Clé	Type	<input checked="" type="checkbox"/> N.	Modèle de date (Ctr...	Length	Precision	Défaut	Comment...
NUMERO_A	<input checked="" type="checkbox"/>	long	<input type="checkbox"/>		10	0		
NOM	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		450	0		
PRENOM	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		450	0		
DOMICILE	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		450	0		

Colonne	Clé	Type	<input checked="" type="checkbox"/> N.	Modèle de date (Ctr...	Length	Precision	Défaut	Comment...
NUMERO_A	<input checked="" type="checkbox"/>	long	<input type="checkbox"/>		10	0		
NOM	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		450	0		
PRENOM	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		450	0		
DOMICILE	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		450	0		

# ATELIERS SUR TALEND OPEN STUDIO

- Modification de ce job :
  - Ouvrir à nouveau le tMap
  - Créer le champ var1 dans la partie droite.
  - Cocher que celui-ci peut être null et qu'il peut avoir une taille de 200 (cf. partie en bas à droite).

The screenshot shows the Talend Open Studio interface for a tMap job. The 'row1' panel on the left lists the input columns: NUMERO\_A, NOM, PRENOM, and DOMICILE. The central 'Var' panel shows a variable named 'var1' with the expression '000001\* Auteur récent\* Auteur ancien' and type 'String'. The 'out' panel on the right lists the output columns: NUMERO\_A, NOM, PRENOM, DOMICILE, and var1. Below the main interface, there are two schema comparison tables for 'ow1' and 'out'.

ow1	Clé	Type	<input checked="" type="checkbox"/>	N.	Modèle de date (Ctr...	Length	Precision	Défaut	Comment...
NUMERO_A	<input checked="" type="checkbox"/>	long	<input checked="" type="checkbox"/>			10	0		
NOM	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>			450	0		
PRENOM	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>			450	0		
DOMICILE	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>			450	0		

out	Clé	Type	<input checked="" type="checkbox"/>	N.	Modèle de date (Ctr...	Length	Precision	Défaut	Comment...
NUMERO_A	<input checked="" type="checkbox"/>	long	<input checked="" type="checkbox"/>			10	0		
NOM	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>			450	0		
PRENOM	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>			450	0		
DOMICILE	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>			450	0		
var1	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>			200			

# ATELIERS SUR TALEND OPEN STUDIO

- Créer un nouveau champ dans la colonne du milieu puis cliquer sur les « ... » pour éditer l'expression :
  - Ajouter le test ternaire dans le nouveau champ :  
(row1.NUMERO\_A>100000)?"Auteur récent":"Auteur ancien"

The screenshot displays the Talend Open Studio interface. On the left, a component named 'row1' is shown with columns: NUMERO\_A, NOM, PRENOM, and DOMICILE. In the center, a 'Var' component is configured with the expression: `00000)?"Auteur récent":"Auteur ancien"`. On the right, an 'out' component is shown with columns: NUMERO\_A, NOM, PRENOM, DOMICILE, and Var.var1. A dialog box titled 'Constructeur d'expression' is open in the foreground, showing the expression editor with the text: `(row1.NUMERO_A>100000)?"Auteur récent":"Auteur ancien"`. The dialog also includes a 'Test' section with a variable list and an 'Aide' section with the instruction: 'Sélectionnez une catégorie et une fonction.'

# ATELIERS SUR TALEND OPEN STUDIO

- Faire glisser le nouveau champ de la colonne du milieu vers la colonne de droite

The screenshot displays the Talend Open Studio interface with two panels side-by-side. The left panel, titled 'Var', contains a table with the following data:

Expression	Type	Variable
00000)?"Auteur récent": "Auteur ancien" ...	String	<input type="checkbox"/> var1

The right panel, titled 'out', contains a table with the following data:

Expression	Column
row1.NUMERO_A	NUMERO_A
row1.NOM	NOM
row1.PRENOM	PRENOM
row1.DOMICILE	DOMICILE
Var.var1	var1

A yellow arrow points from the 'var1' variable in the 'Var' panel to the 'Var.var1' row in the 'out' panel, illustrating the mapping process.

# ATELIERS SUR TALEND OPEN STUDIO

- Exécuter à nouveau le job, vous devriez obtenir ce type de résultat

The screenshot shows the Talend Open Studio interface. At the top, there are three tabs: 'Job test 0.1', 'Job LectureFichiertoBDD 0.1', and '\*Job lectureBDD 0.1'. The main workspace displays a job design with three components: 'auteur', 'tMap\_1', and 'tLogRow\_1'. The 'auteur' component is connected to 'tMap\_1', which is connected to 'tLogRow\_1'. The 'tLogRow\_1' component shows a preview of the data output.

Job lectureBDD

Exécution

Exécuter Arrêter Effacer

Démarrage du job lectureBDD à 02:12 10/02/2020.  
[statistics] connecting to socket on port 3747  
[statistics] connected

NUMERO_A	NOM	PRENOM	DOMICILE	var1
8547	Chambord	Emilie	Nice	Auteur ancien
52136	Fabiere	Sylvie	Bordeaux	Auteur ancien
78545	TINTIN	Thierry	Clermont	Auteur ancien
85478	Castafiore	Emilie	Paris	Auteur ancien
542536	Dupont	Pierre	Avignon	Auteur récent
8547585	Momo	Roland	Toulouse	Auteur récent

[statistics] disconnected  
Le Job lectureBDD s'est terminé à {1,date,HH:mm dd/MM/yyyy}. [code de sortie={2,number}]

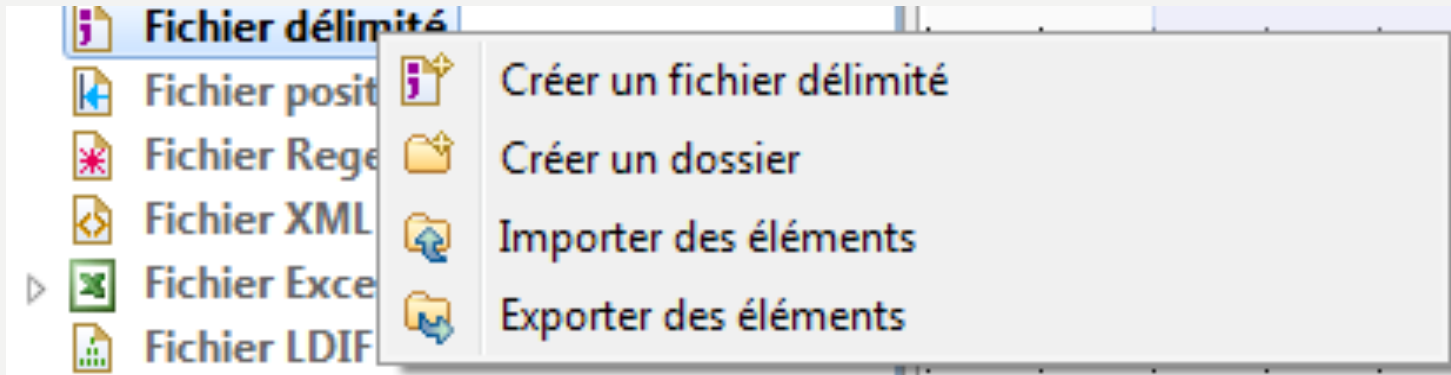
# TO DO

- 2<sup>ème</sup> job pour lire un fichier CSV et enregistrer les données en BDD
  - Création d'un deuxième job
  - Lecture d'un fichier CSV
  - Création d'une routine
  - Ajout d'un composant tMap
  - Création d'un nouveau champ
  - Ajout d'un composant pour écrire dans une BDD
  - Exécution du job



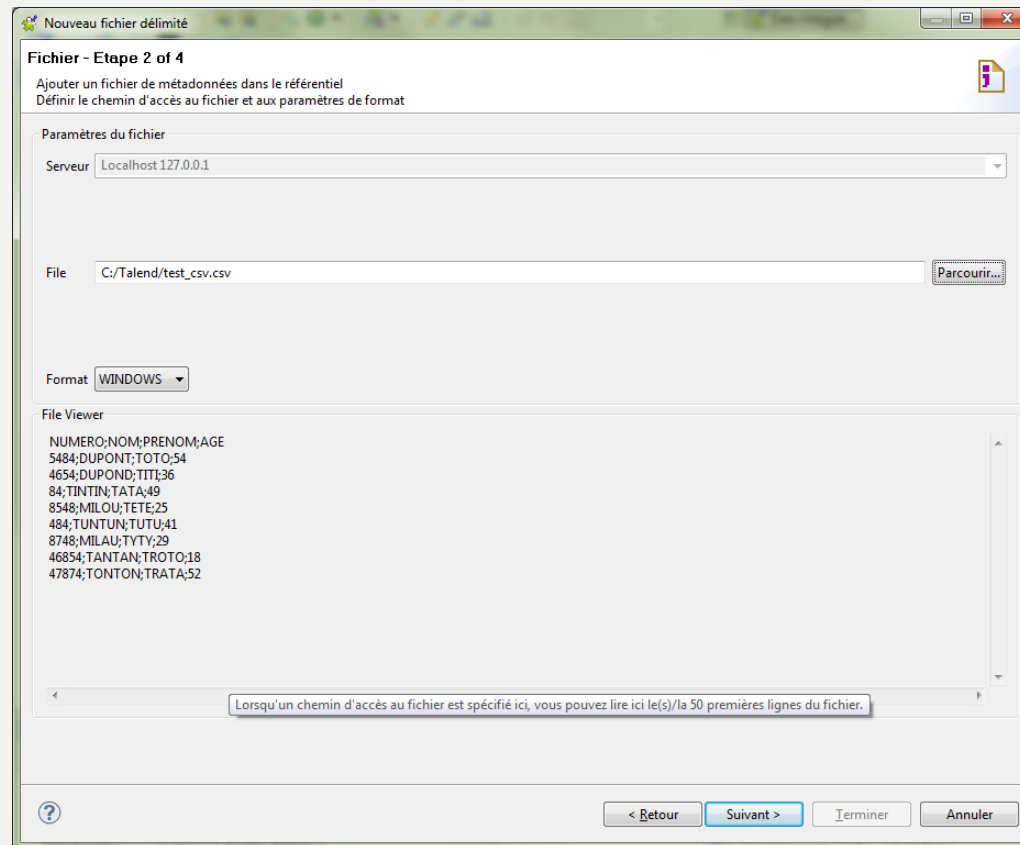
# ATELIERS SUR TALEND OPEN STUDIO

- Création d'une connexion à un fichier CSV



# ATELIERS SUR TALEND OPEN STUDIO

- Création d'une connexion à un fichier CSV
  - Lire le fichier test\_CSV.csv



# ATELIERS SUR TALEND OPEN STUDIO

- Création d'une connexion à un fichier CSV
  - Définir les lignes d'en-tête comme nom de colonnes

Editer un fichier délimité existant

**Fichier - Etape 3 de 3**  
Update an existing Metadata File on repository  
Define the setting of the parse job

Paramètres de fichier

Encodage: US-ASCII

Séparateur de champs: Semicolon | Caractère correspondant: ";"

Séparateur de lignes: Standard EOL | Caractère correspondant: "\n"

Paramètres du caractère d'échappement

CSV  Délimité

Caractère d'échappement: Vide

Entourage du texte: Vide

Découper la ligne avant le champ

Lignes à ignorer

Si des lignes doivent être ignorées, spécifiez les paramètres suivants

En-tête:  1

Pied de page:

Ignorer les lignes vides

Limite de lignes

Si le nombre de lignes doit être limité, spécifiez ce nombre.

Limite:

Aperçu | Sortie

Définir les lignes d'en-tête comme nom de colonnes [Rafraichir l'aperçu](#)

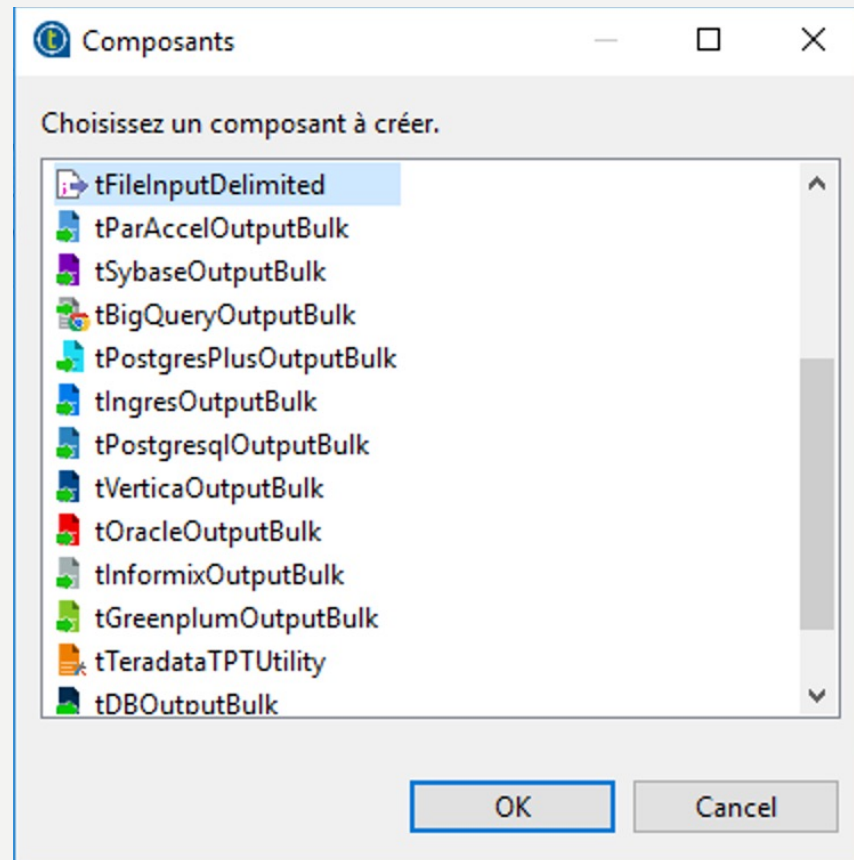
NUMERO	NOM	PRENOM	AGE
5484	DUPONT	TOTO	54
4654	DUPOND	TITI	36
84	TINTIN	TATA	49
8548	MILOU	TETE	25
484	TUNTUN	TUTU	41
8748	MILAU	TYTY	29

[Exporter en tant que contexte](#) [Revenir au contexte précédent](#)

< Back Next > **Finish** Cancel

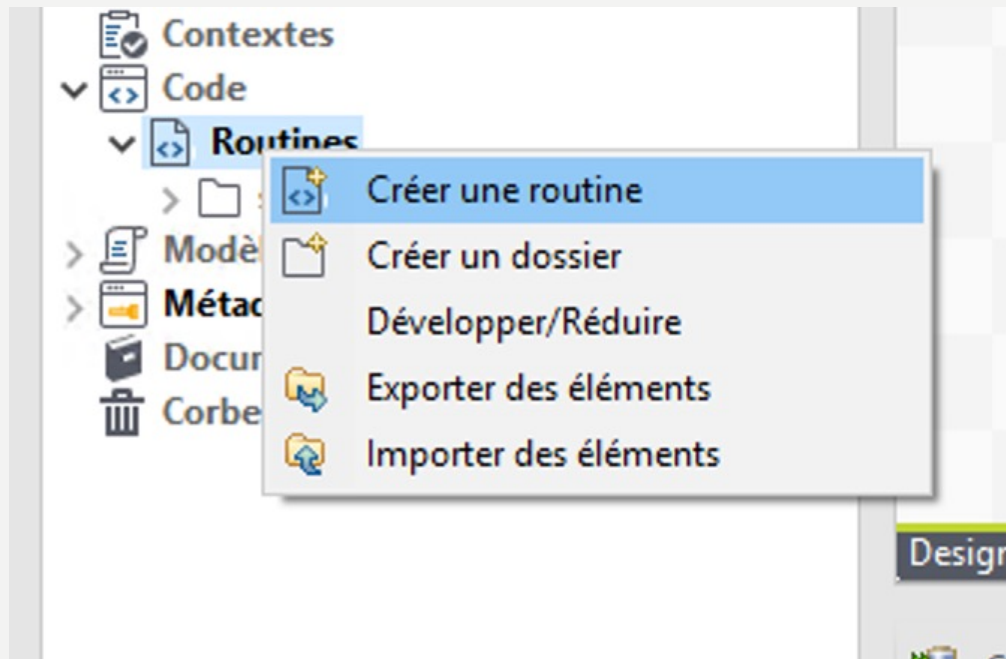
# ATELIERS SUR TALEND OPEN STUDIO

- Glisser la connexion au fichier CSV dans la partie Designer, Talend proposera la lecture via le composant « tFileInputDelimited »



# ATELIERS SUR TALEND OPEN STUDIO

- Création d'une routine :
  - Une routine permet d'exécuter du code spécifique



# ATELIERS SUR TALEND OPEN STUDIO

- Création d'une routine :

```
public static String Concat (String message1,  
String message2)  
{  
    String resultat = new String("");  
    resultat = message1 + " / " + message2;  
    return resultat;  
}
```

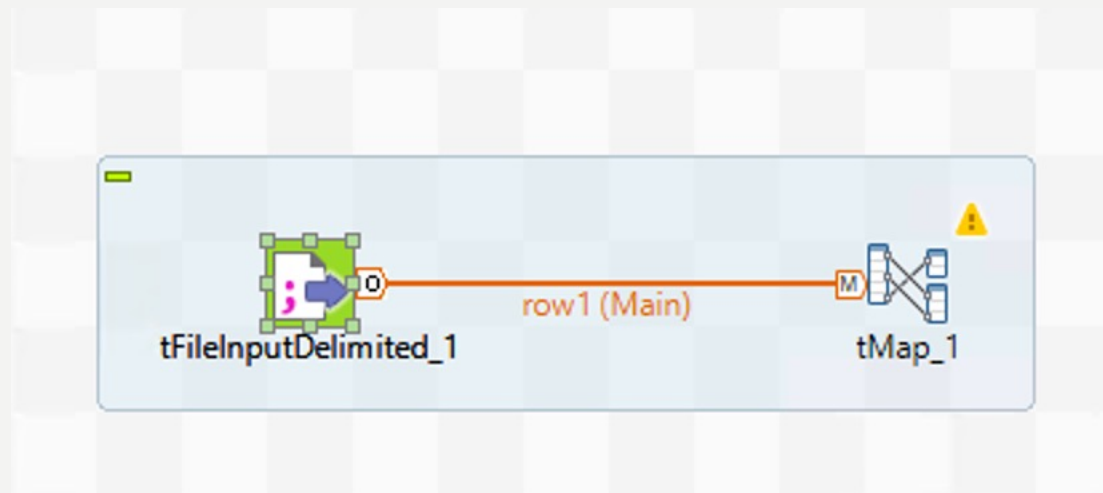
# ATELIERS SUR TALEND OPEN STUDIO

- Création d'une routine :

```
public class testRoutine {  
    /**  
     * helloExample: not return value, only print "hello" + message.  
     *  
     *  
     * {talendTypes} String  
     *  
     * {Category} User Defined  
     *  
     * {param} string("world") input: The string need to be printed.  
     *  
     * {example} helloExemple("world") # hello world !.  
     */  
    public static void helloExample(String message) {  
        if (message == null) {  
            message = "World"; //$NON-NLS-1$  
        }  
        System.out.println("Hello " + message + " !"); //$NON-NLS-1$ //$NON-NLS-2$  
    }  
  
    public static String Concat (String message1, String message2)  
    {  
        String resultat = new String("");  
        resultat = message1 + " / " + message2;  
        return resultat;  
    }  
}
```

# ATELIERS SUR TALEND OPEN STUDIO

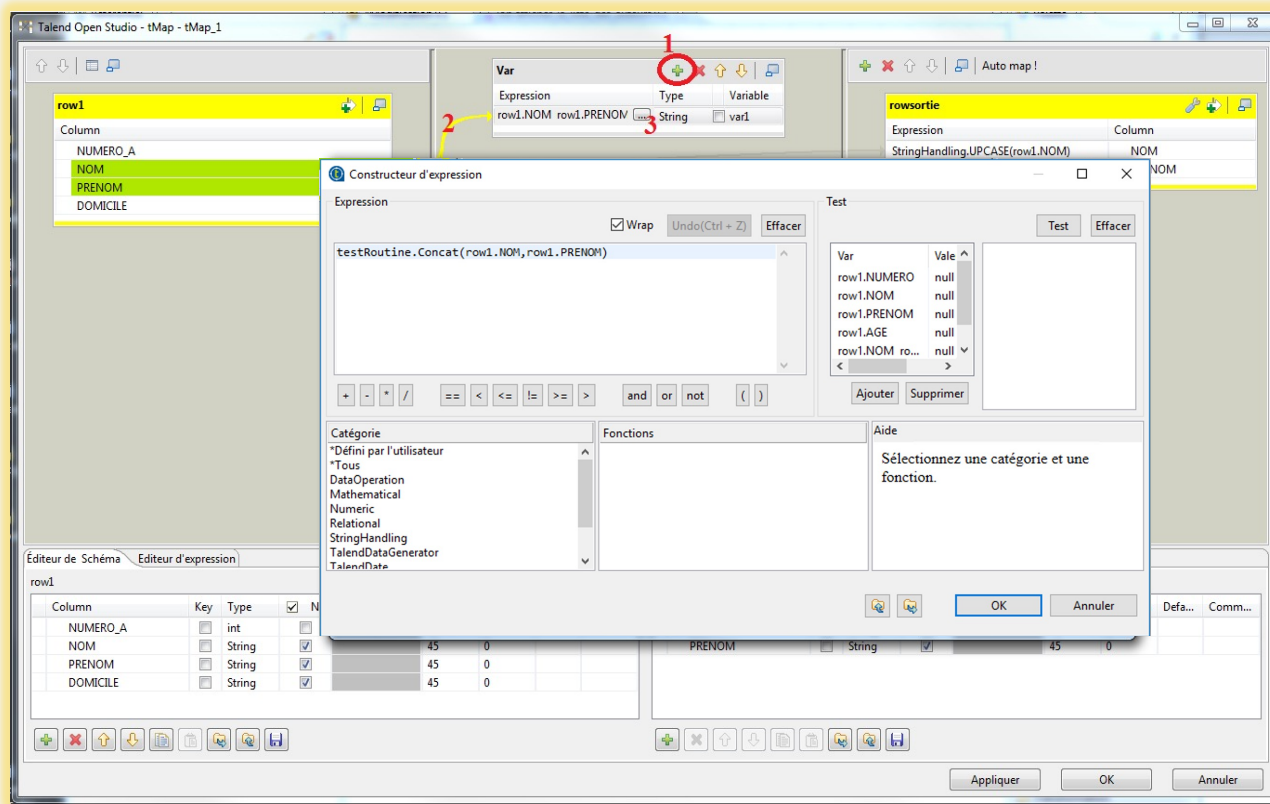
- Ajout du composant tMap à la suite du composant « tFileInputDelimited »





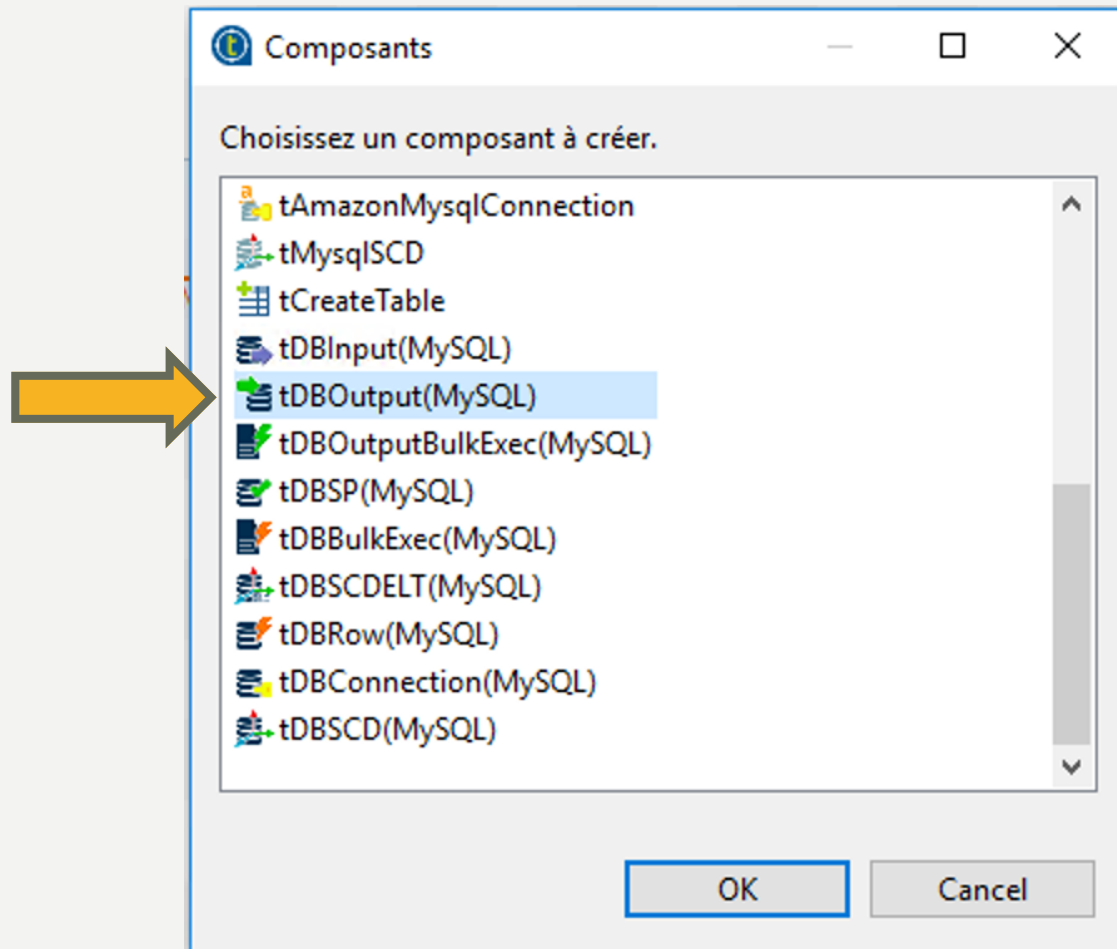
# ATELIERS SUR TALEND OPEN STUDIO

- Création d'une nouveau champ (cf. ci-dessous)
- Puis modification de l'expression :  
testRoutine.Concat(row1.NOM,row1.PRENOM)



# ATELIERS SUR TALEND OPEN STUDIO

- Chargement dans une base :



# ATELIERS SUR TALEND OPEN STUDIO

- Chargement dans une base :
  - Vérifier que les mêmes paramètres soient bien positionnés
  - Dans un premier temps, mettre l'option Créer la table si elle n'existe pas dans le champ « Table »
  - Puis après la première exécution, mettre comme sur la capture ci-dessous :

The screenshot shows the configuration window for a Talend connector named "tDBOutput\_1(MySQL)". The window is divided into several sections:

- Paramètres simples:** Database is set to "MySQL" with an "Apply" button.
- Paramètres avancés:** Type de propriété is "Référentiel", and Bases de données (MYSQL):connexionBDDM is selected.
- Paramètres dynamiques:** Version de la base de données is "MySQL 5".
- View:**  Utiliser une connexion existante.
- Documentation:** Hôte is "localhost", Port is "3306", Base de données is "gestion\_livres", Utilisateur is "root", and Mot de passe is "\*\*\*\*\*".
- Table:** The table name is "insertCSVFile".
- Action sur la table:** "Supprimer la table si elle existe et la créer".
- Action sur les données:** "Insert".
- Schéma:** "Built-In" with buttons for "Modifier le schéma" and "Sync colonnes".
- Source de données:** A section with a note "This option only applies when deploying and running in the Talend Runtime" and a checkbox for "Spécifier l'alias de la source de données".
- Arrêter en cas d'erreur.

# ATELIERS SUR TALEND OPEN STUDIO

- Liaison avec le tMap entre les composants d'entrée et de sortie
  - Vérifier les paramètres
  - Rajouter le nouveau champ dans la sortie

The screenshot displays the Talend Open Studio interface for configuring a tMap component. The main workspace is divided into three panes: the input table 'row1', the variable definition table 'Var', and the output table 'out'. The 'row1' table lists columns: NUMERO, NOM, PRENOM, and AGE. The 'Var' table shows a variable 'var1' of type 'String' with the expression 'testRoutine.Concat(row1.NOM,row1.PRE...'. The 'out' table lists columns: NUMERO, NOM, PRENOM, AGE, and var1. A yellow arrow points from the 'var1' variable to the 'var1' column in the 'out' table. Below the main workspace, there are two schema editor panels. The left panel, titled 'row1', shows a table with columns: Colonne, Clé, Type, N., Modèle de date (Ctr..., Length, Precision, Défaut, and Comment... The right panel, titled 'out', shows a table with columns: Colonne, Clé, Type, N., Modèle de date (Ctr..., Length, Precision, Défaut, and Comment... The bottom of the window features a toolbar with various icons and buttons for 'Appliquer', 'OK', and 'Annuler'.

Colonne	Clé	Type	<input checked="" type="checkbox"/> N.	Modèle de date (Ctr...	Length	Precision	Défaut	Comment...
NUMERO	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		5	0		
NOM	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		6	0		
PRENOM	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		5	0		
AGE	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		2	0		

Colonne	Clé	Type	<input checked="" type="checkbox"/> N.	Modèle de date (Ctr...	Length	Precision	Défaut	Comment...
NUMERO	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		5	0		
NOM	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		6	0		
PRENOM	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		5	0		
AGE	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		2	0		
var1	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		50			

# ATELIERS SUR TALEND OPEN STUDIO

- Exécution du job
- Vérifier dans la BDD

The screenshot shows a job execution progress bar with the following details:

- tFileInputDelimited\_1: 9 rows in 0,02s, 562,5 rows/s, row1 (Main)
- tMap\_1: 9 rows in 0,45s, 19,87 rows/s, out (Main)
- tDBOutput\_1: (no statistics shown)

Below the progress bar, the execution console shows the following log:

```
Démarrage du job LectureFichiertoBDD a 03:01 10/02/2020.  
[statistics] connecting to socket on port 3753  
[statistics] connected  
[statistics] disconnected  
  
Le Job LectureFichiertoBDD sest terminé à {1,date,HH:mm dd/MM/yyyy}. [code de sortie={2,number}]
```

The screenshot shows a database schema tree on the left and a result grid on the right.

**Database Schema Tree:**

- gestion\_livres
  - Tables
    - auteur
    - insertcsvfile
    - insertgoogle
  - Views
  - Stored Procedures
  - Functions
- sakila
- sys
- world

**Result Grid:**

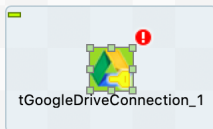
NUMERO	NOM	PRENOM	AGE	null
5484	DUPONT	TOTO	54	DUPONT / TOTO
4654	DUPOND	TITI	36	DUPOND / TITI
84	TINTIN	TATA	49	TINTIN / TATA
8548	MILOU	TETE	25	MILOU / TETE
484	TUNTUN	TUTU	41	TUNTUN / TUTU
8748	MILAU	TYTY	29	MILAU / TYTY
46854	TANTAN	TROTO	18	TANTAN / TROTO
47874	TONTON	TRATA	52	TONTON / TRATA

# TO DO

- 3<sup>ème</sup> job : récupération d'un fichier CSV sur Google Drive et écriture du fichier CSV sur le disque en local
  - Ajout du composant `GoogleDriveConnection`
  - Installation des trois bibliothèques
  - Configuration du composant `GoogleDriveConnection`
  - Ajout du composant `tGoogleDriveGet` et configuration
  - Exécution du job avec validation via le navigateur pour s'authentifier correctement
  - Ajout d'un composant `tFileOutputDelimited`
  - Exécution du job
  - Comparaison des fichiers
  - Lien possible avec le job 2

# ATELIERS SUR TALEND OPEN STUDIO

- Connexion à Google Drive
  - Glisser le composant suivant dans un nouveau job
  - Cliquer sur « installer »



.1) Contexts(test) Composant Exécuter (Job test)

**DriveConnection\_1**

**mples**

**ngs**

**amiques**

**l**

Use Proxy

**Warning:** Ce composant tGoogleDriveConnection requiert l'installation d'au moins un Jar externe. [Installer...](#)

Type de propriété: Built-in Application Name: ""

OAuth Method: Installed Application (Id & Secret)

Client ID: "" Client Secret: \*\*\*\*

# ATELIERS SUR TALEND OPEN STUDIO

- Télécharger et installer les trois bibliothèques les unes après les autres

Le composant tGoogleDriveConnection requiert les modules tiers suivants :

Liste des modules non installés pour le composant tGoogleDriveConnection

Jar	Requis par composant	Requis	Licence	Maven URI	Plus d'informations	Action
httpclient.jar	tGoogleDriveConnection	<input checked="" type="checkbox"/>	Apache-2.0	mvn:org.apache.httpcomponents/http...		Télécharger et installer
httpcore.jar	tGoogleDriveConnection	<input checked="" type="checkbox"/>	Apache-2.0	mvn:org.apache.httpcomponents/http...	<a href="http://hc.apache.org/">http://hc.apache.org/</a>	Télécharger et installer
slf4j-log4j12-1.7.2.jar	tGoogleDriveConnection	<input checked="" type="checkbox"/>	X11	mvn:org.talend.libraries/slf4j-log4j12...	<a href="http://www.slf4j.org">http://www.slf4j.org</a>	Télécharger et installer

Ne plus afficher

Close      Télécharger et installer tous les modules disponibles.



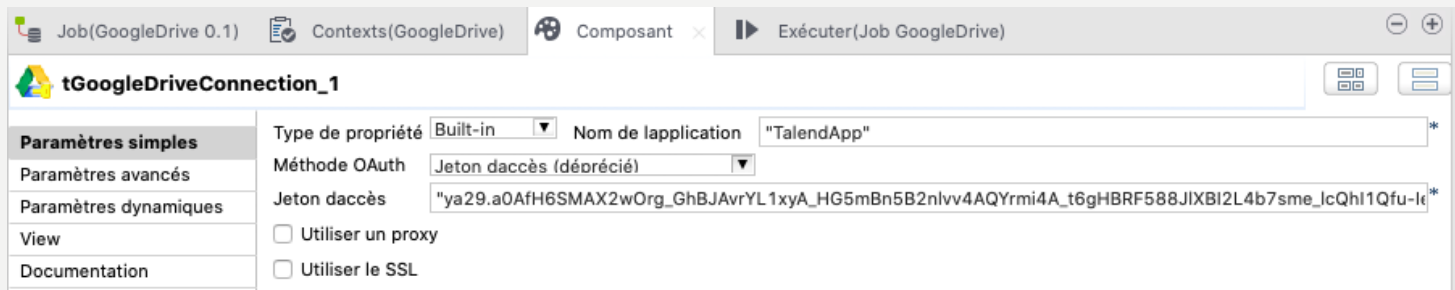
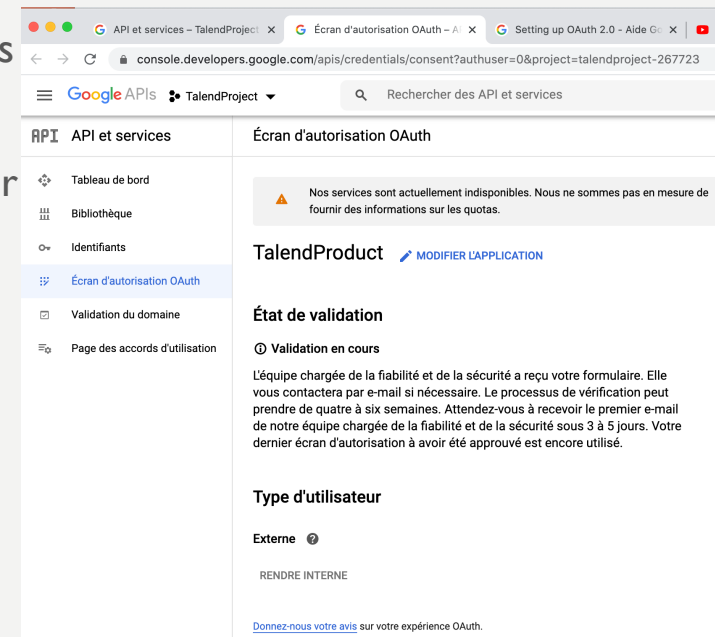
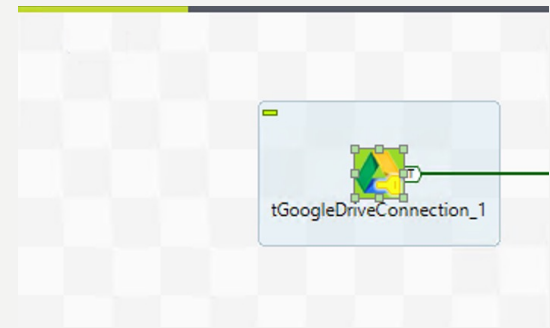
# TO DO

- Connexion à Google Drive

Auparavant, nous pouvions utiliser une connexion via OAuth mais dernièrement Google a renforcé son niveau de sécurité pour l'accès à des données sur les drives.

Nous allons utiliser une méthode plus manuelle pour obtenir un token temporairement.

- Rentrer les paramètres suivants :
  - Nom de l'application : « TalendApp »
  - Jeton d'accès : cf. page suivante



# TO DO

- Pour obtenir un token, ouvrir la page suivante :

<https://developers.google.com/oauthplayground/>





- Puis cocher sur Drive API v3 et enfin sur la première API :

<https://www.googleapis.com/authdrive>

- Puis cliquer sur Authorize APIs
- Vous allez être redirigé vers une page d'authentification :
  - Adresse mail : [tpisimaetl@gmail.com](mailto:tpisimaetl@gmail.com)
  - Mot de passe : recopier le tableau

Step 1 Select & authorize APIs

Select the scope for the APIs you would like to access or input your own OAuth scopes below. Then click the "Authorize APIs" button.

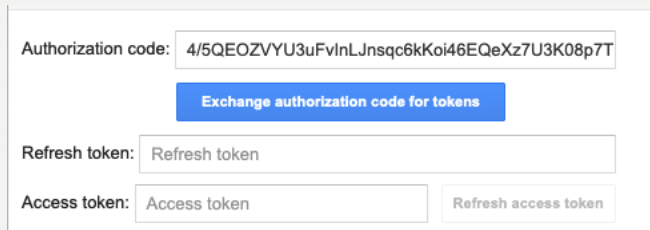
- ▶  Dialogflow API v2
- ▶  Display & Video 360 API v1
- ▶  DoubleClick Bid Manager API v1.1
- ▼  Drive API v3
  - ✓ <https://www.googleapis.com/auth/drive>
  - <https://www.googleapis.com/auth/drive.appdata>
  - <https://www.googleapis.com/auth/drive.file>
  - <https://www.googleapis.com/auth/drive.metadata>
  - <https://www.googleapis.com/auth/drive.metadata.readonly>
  - <https://www.googleapis.com/auth/drive.photos.readonly>
  - <https://www.googleapis.com/auth/drive.readonly>
  - <https://www.googleapis.com/auth/drive.scripts>

Input your own scopes

Authorize APIs

# TO DO

- Appuyer sur autoriser
- Puis cliquer sur le bouton « Exchange Authorization code for tokens »



- Récupérer l'access\_token et recopier le dans le champs jeton d'accès du composant Talend

```
{
  "access_token":
  "ya29.a0AfH63CM1ZUYDe2l8s_CoHYVOCkDANrBzEwFLaB9C-WH4N4128ThumapOubabMYPDIeO1uesT0annidh5NEd4baDHe8aGcWY2Dcu13CFs
  "scope": "https://www.googleapis.com/auth/drive",
  "token_type": "Bearer",
  "expires_in": 3599,
  "refresh_token": "1//04GpPgwmHNdf3CgYIARAAGAQSfwF-
```



# ATELIERS SUR TALEND OPEN STUDIO

- Glisser le composant suivant dans un nouveau job
  - tGoogleDriveGet\_1



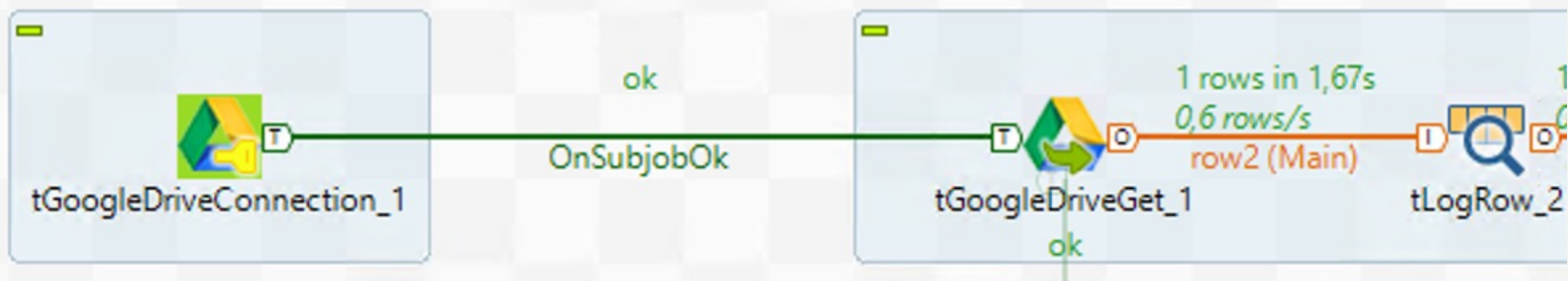
- Rentrer les paramètres suivants :

tGoogleDriveGet_1	
Paramètres simples	Composant de connexion <input type="text" value="tGoogleDriveConnection_1"/>
Paramètres avancés	Fichier <input type="text" value="TP/test_csv.csv"/> * Méthode d'accès <input type="text" value="par Nom"/> *
Paramètres dynamiques	<input checked="" type="checkbox"/> Sauvegarder en tant que fichier
View	Sauvegarder dans <input type="text" value="C:/Users/jofontan/Downloads/test_OUTPUT_GET.csv"/> *
Documentation	Schéma <input type="text" value="Built-in"/> Modifier le schéma ...

- Fichier : TP/test\_csv.csv
- Sauvegarder dans : lien où vous souhaitez que le fichier soit déposé

# ATELIERS SUR TALEND OPEN STUDIO

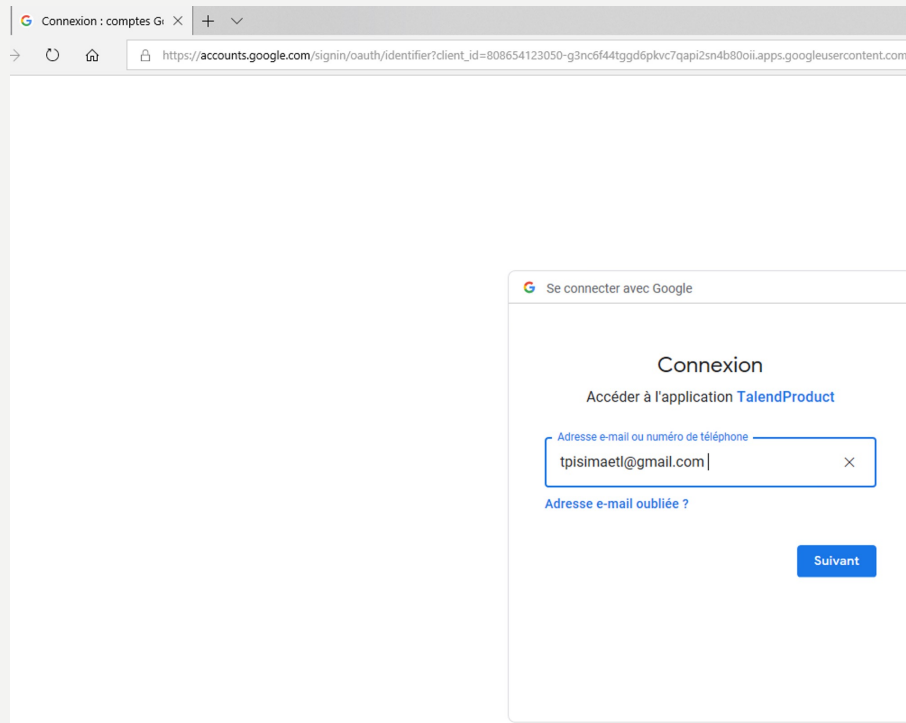
- Rajouter un tLogRow à la suite du composant tGoogleDriveGet\_I pour lire le résultat
- Connecter les composants avec les conditions suivantes (bouton droit « trigger »)



Job test	
Exécution simple	Exécution
Exécution Debug	<input type="button" value="▶ Exécuter"/> <input type="button" value="■ Arrêter"/> <input type="button" value="🗑 Effacer"/>
Paramètres avancés	
Target Exec	
Memory Run	
	<pre>Démarrage du job test a 03:19 10/02/2020. [statistics] connecting to socket on port 3337 [statistics] connected NUMERO; NOM; PRENOM; AGE 5484; DUPONT; TOTO; 54 4654; DUPOND; TITI; 36 84; TINTIN; TATA; 49 8548; MILOU; TETE; 25 484; TUNTUN; TUTU; 41 8748; MILAU; TYTY; 29 46854; TANTAN; TROTO; 18 47874; TONTON; TRATA; 52  [statistics] disconnected  Le Job test sest terminé à {1,date,HH:mm dd/MM/yyyy}. [code de sortie={2,number}]</pre>

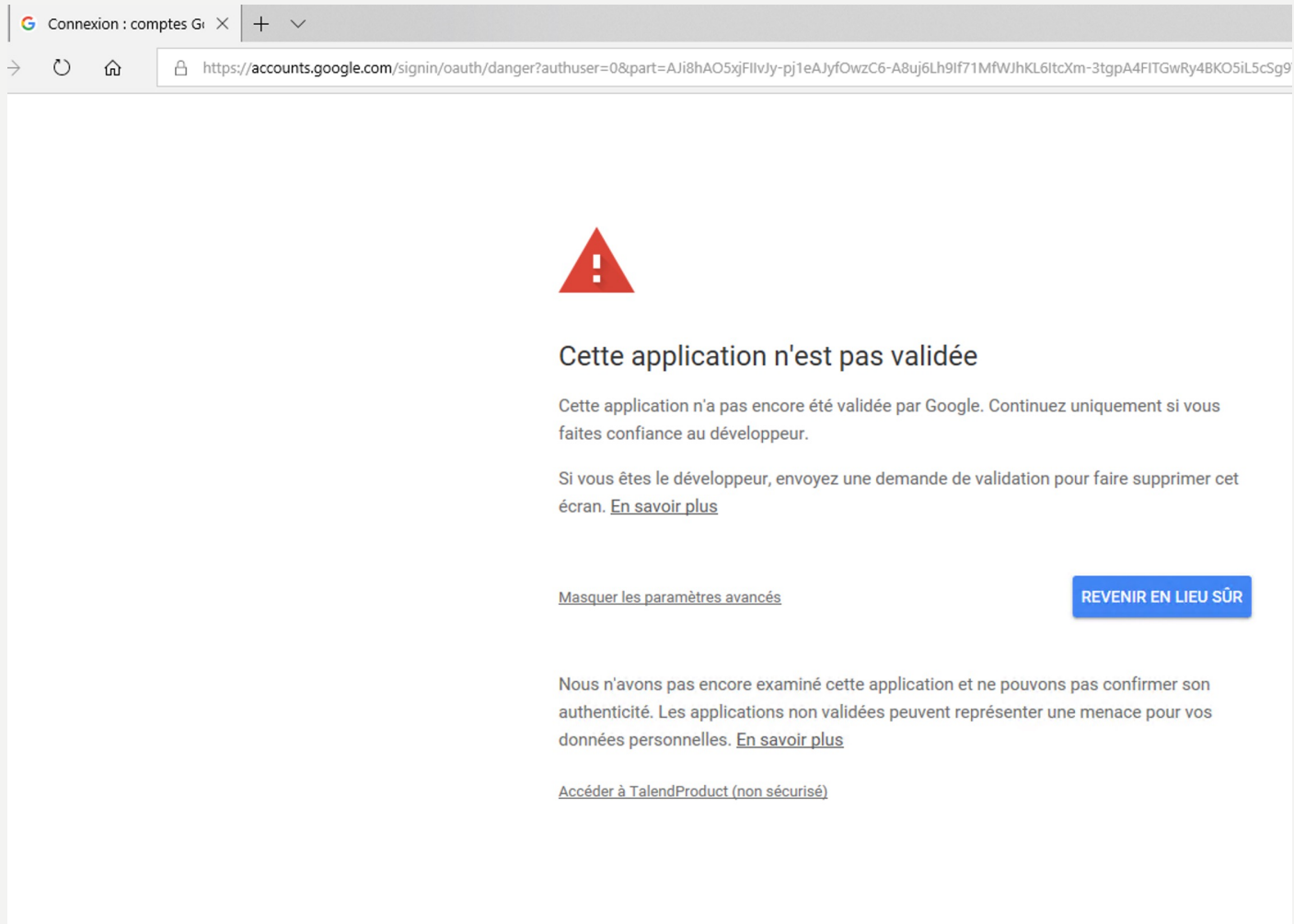
# ATELIERS SUR TALEND OPEN STUDIO

- Exécuter le job
- Il se peut qu'il vous demande de vous authentifier
  - Adresse mail : [tpisimaetl@gmail.com](mailto:tpisimaetl@gmail.com)
  - Mot de passe : au tableau




# ATELIERS SUR TALEND OPEN STUDIO

Cliquer sur les paramètres avancés :



Connexion : comptes G × + ▾

→ ↻ 🏠 <https://accounts.google.com/signin/oauth/danger?authuser=0&part=AJi8hAO5xjFIlvJy-pj1eAJyfOwzC6-A8uj6Lh9If71MfWJhKL6ItcXm-3tgpA4FITGwRy4BKO5iL5cSg9>



**Cette application n'est pas validée**

Cette application n'a pas encore été validée par Google. Continuez uniquement si vous faites confiance au développeur.

Si vous êtes le développeur, envoyez une demande de validation pour faire supprimer cet écran. [En savoir plus](#)

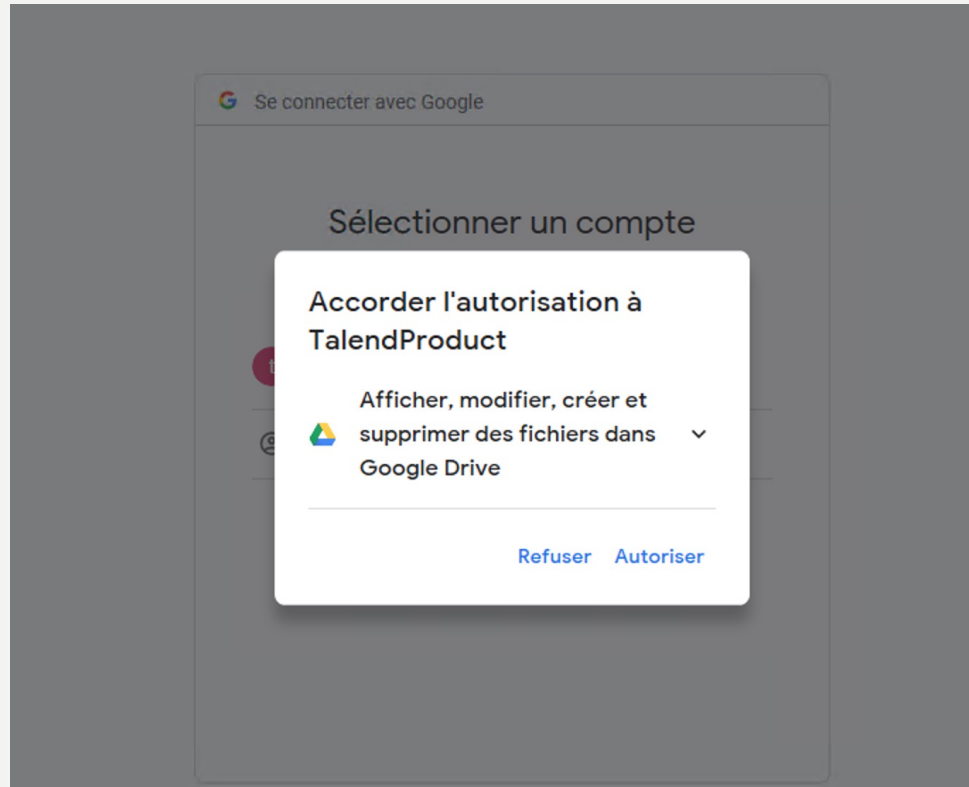
[Masquer les paramètres avancés](#) [REVENIR EN LIEU SÛR](#)

Nous n'avons pas encore examiné cette application et ne pouvons pas confirmer son authenticité. Les applications non validées peuvent représenter une menace pour vos données personnelles. [En savoir plus](#)

[Accéder à TalendProduct \(non sécurisé\)](#)

# ATELIERS SUR TALEND OPEN STUDIO

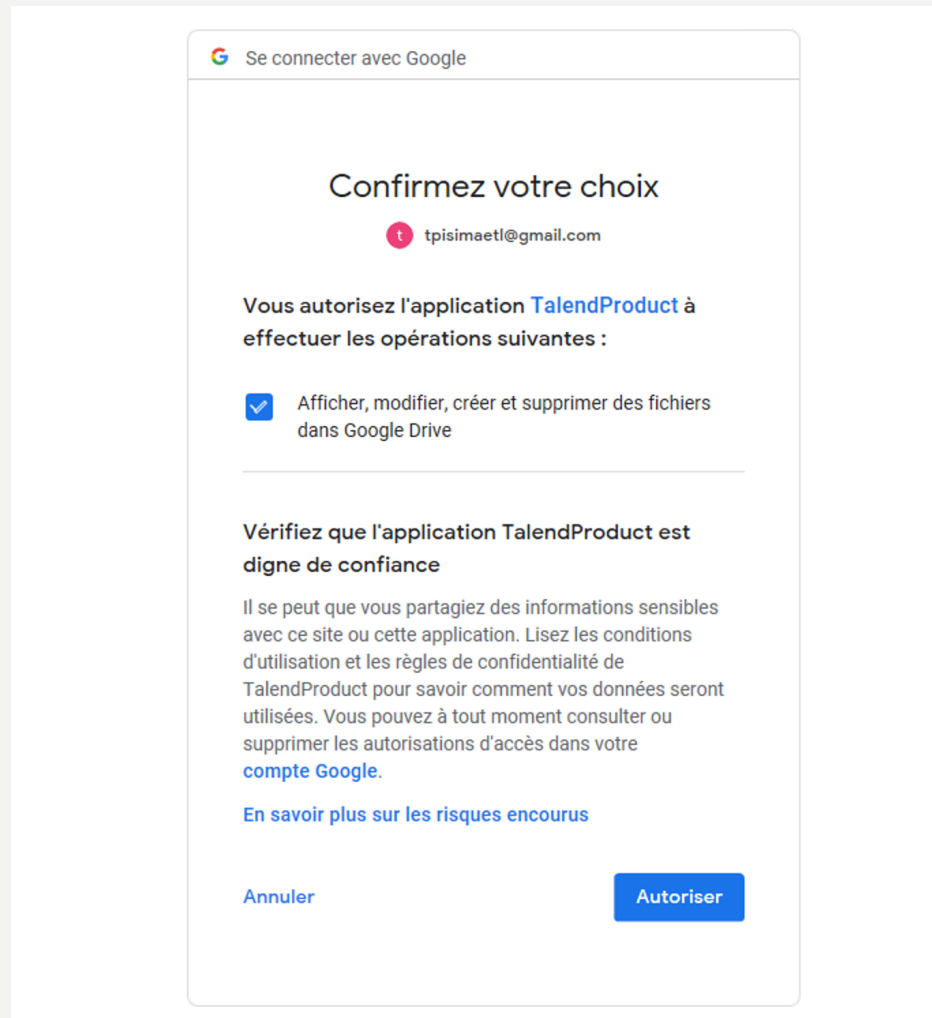
- Autoriser TalendProduct à accéder à GoogleDrive





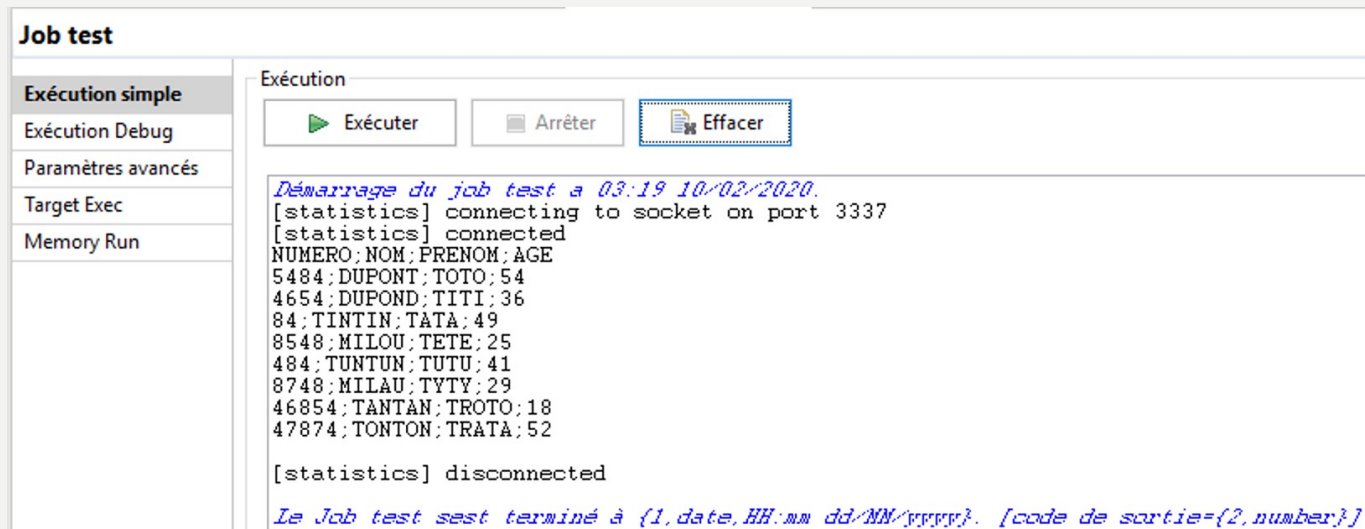
# ATELIERS SUR TALEND OPEN STUDIO

- Autoriser TalendProduct à accéder à GoogleDrive et confirmer votre choix



# ATELIERS SUR TALEND OPEN STUDIO

- Trace de l'exécution



**Job test**

Exécution

▶ Exécuter    ⏹ Arrêter    🗑 Effacer

```
Démarrage du job test a 03:19 10/02/2020.  
[statistics] connecting to socket on port 3337  
[statistics] connected  
NUMERO;NOM;PRENOM;AGE  
5484;DUPONT;TOTO;54  
4654;DUPOND;TITI;36  
84;TINTIN;TATA;49  
8548;MILOU;TETE;25  
484;TUNTUN;TUTU;41  
8748;MILAU;TYTY;29  
46854;TANTAN;TROTO;18  
47874;TONTON;TRATA;52  
  
[statistics] disconnected  
  
Le Job test s'est terminé à {1,date,HH:mm dd/MM/yyyy}. [code de sortie={2,number}]
```

- Vous devriez avoir deux fichiers CSV strictement identiques.
- Si les répertoires entre le job 2 et le job 3 correspondent, vous avez créé une chaîne d'intégration

# ATELIERS SUR TALEND OPEN STUDIO

- Pour aller plus loin
  - Glisser le composant tGoogleDriveList
  - Configurer ce composant comme la capture ci-dessous :

The screenshot displays the Talend Open Studio Designer interface. At the top, a workflow diagram shows an 'OnSubJobOk' trigger connected to a 'tGoogleDriveList\_1' component, which is then connected to a 'tLogRow\_1' component. Below the canvas, the configuration panel for 'tGoogleDriveList\_1' is visible. The 'Paramètres simples' tab is active, showing the following settings:

Paramètres simples	Configuration
Composant de connexion	tGoogleDriveConnection_1
Paramètres avancés	Nom du dossier: "TP" * Méthode d'accès: par Nom *
Paramètres dynamiques	Type de fichier dans la liste FileList: Files
View	<input checked="" type="checkbox"/> Inclure les sous-répertoires
Documentation	Schéma: Built-in Modifier le schéma ...

- Brancher un tLogRow pour regarder le résultat

# ATELIERS SUR TALEND OPEN STUDIO

- Pour aller plus loin
  - Trace de l'exécution :

The screenshot displays the Talend Open Studio interface during a job execution. The top section shows a job diagram with the following components and their execution statistics:

- tGoogleDriveConnection\_1**: Connected to **tGoogleDriveGet\_1** via an **OnSubjobOk** trigger.
- tGoogleDriveGet\_1**: Executed 1 row in 1,56s at 0,64 rows/s. It is connected to **tLogRow\_2** via a **row2 (Main)** trigger.
- tLogRow\_2**: Executed 1 row in 0,00s at 0,00 rows/s. It is connected to **tFileOutputDelimited\_1** via a **row3 (Main)** trigger.
- tGoogleDriveList\_1**: Executed 2 rows in 0,53s at 3,77 rows/s. It is connected to **tLogRow\_1** via a **row1 (Main)** trigger.

The bottom section shows the 'Job test' execution console. The log contains the following text:

```
Démarrage du job test à 03:31 10/02/2020.
[statistics] connecting to socket on port 3909
[statistics] connected
1_yI_0q8CyZA8vlix72FJI5WCYMB9iONA|testFetch_csv.csv|text/csv|2020-02-10T00:55:19Z|190|drive#file|false|[1DtrUBaad4gWGr_lw328srGmGmbHvshA7]|https://drive.google.com/file/d/1_yI_0q8CyZA8vlix72FJI5WCYMB9iONA/view?usp=drivesdk
1vBuIGat8mxQkLIOnhFpnnZTO93qcR-tL|test_csv.csv|text/csv|2020-02-10T00:27:30Z|190|drive#file|false|[1DtrUBaad4gWGr_lw328srGmGmbHvshA7]|https://drive.google.com/file/d/1vBuIGat8mxQkLIOnhFpnnZTO93qcR-tL/view?usp=drivesdk
[statistics] disconnected

Le Job test s'est terminé à {1,date,HH:mm dd/MM/yyyy}. [code de sortie={2,number}]
```

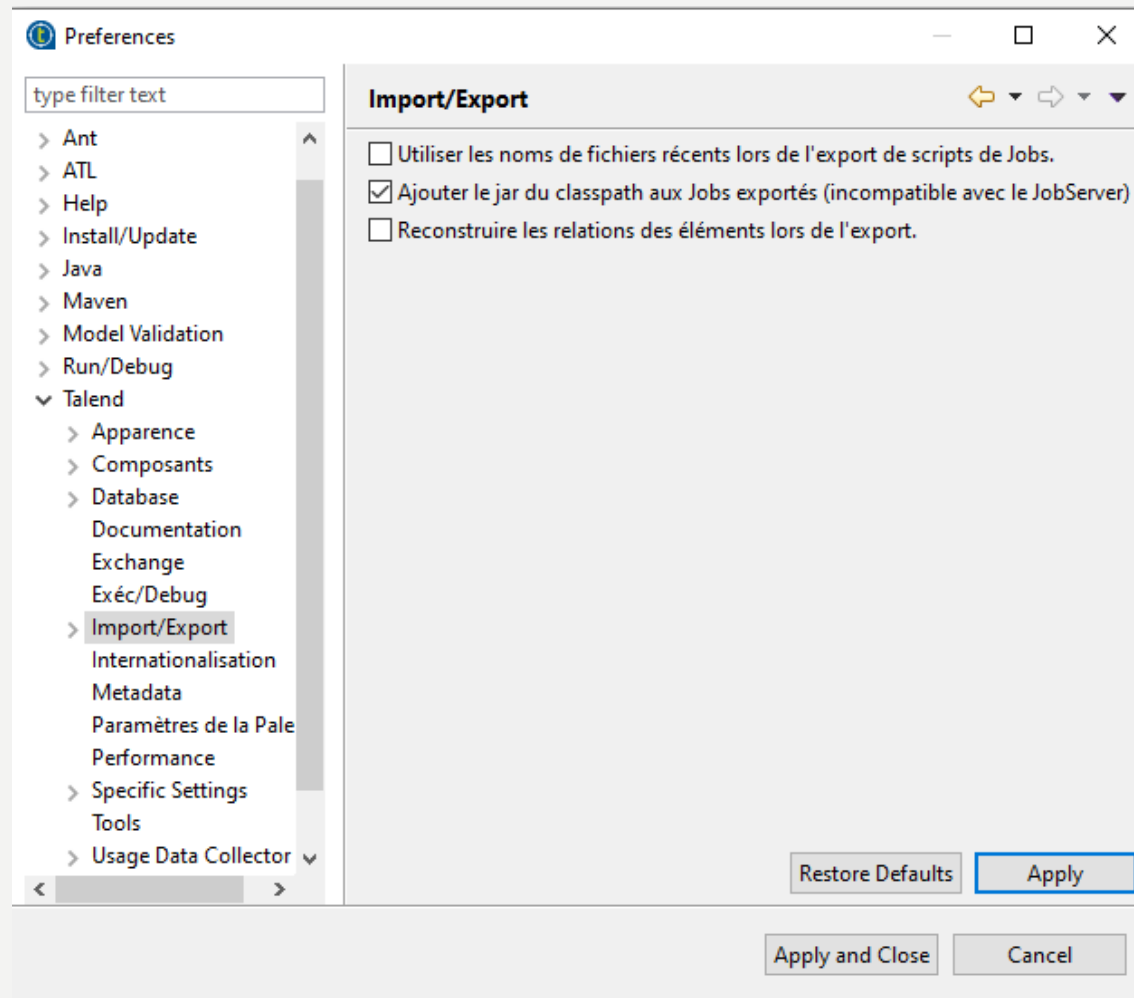
- Que permet de faire ce composant ?

# TO DO

- 4<sup>ème</sup> job : Déploiement d'un job
  - Exporter son job
  - Créer une tâche planifiée qui lance son job
  - Gestion des configurations via les contextes
  - Exécution du job standalone en fonction de l'environnement

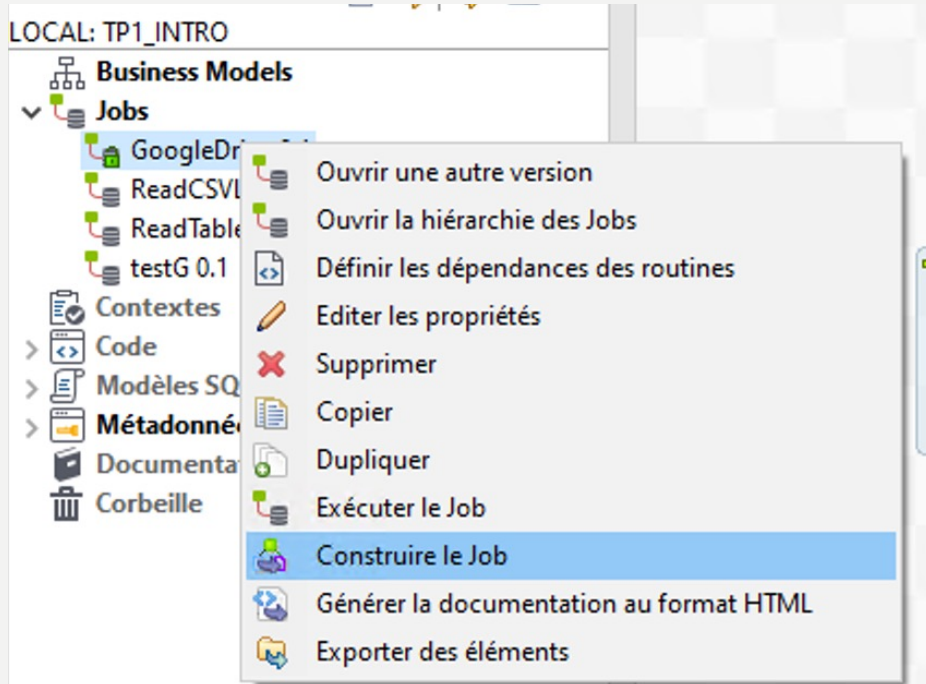
# ATELIERS SUR TALEND OPEN STUDIO

- Exporter le job pour programmer son lancement :
  - Configurer votre studio pour inclure toutes les jars dans un fichier



# ATELIERS SUR TALEND OPEN STUDIO

- Exporter le job pour programmer son lancement :
  - Faire un clic droit sur le job que vous voulez construire :



# ATELIERS SUR TALEND OPEN STUDIO

- Exporter le job pour programmer son lancement :
  - Positionner les mêmes paramètres que ci-dessous

Construire le Job

Vers le fichier archive : C:\Talend\TOS\_DI-Win32-20200219\_1130-V7.3.1\GoogleDrive\_0.1.zip Browse...

Job Version  
Sélectionnez la version de Job 0.1

Type de construction  
Sélectionner le type de construction Job standalone  Extraire le fichier zip

Options  
 Interpréteur de commande Tout  
 Scripts de contexte Default  Appliquer le contexte aux Jobs fils  
 Écraser les valeurs des paramètres

>>  
 Eléments  
 Sources Java



# ATELIERS SUR TALEND OPEN STUDIO

- Exporter le job pour programmer son lancement :
  - Exécuter le .bat via invite de commande

```
Invite de commandes
Microsoft Windows [version 10.0.19041.572]
(c) 2020 Microsoft Corporation. Tous droits réservés.

C:\Users\jofontan>cd C:\Users\jofontan\Desktop\GoogleDrive_0.1\GoogleDrive

C:\Users\jofontan\Desktop\GoogleDrive_0.1\GoogleDrive>GoogleDrive_run.bat

C:\Users\jofontan\Desktop\GoogleDrive_0.1\GoogleDrive>C:

C:\Users\jofontan\Desktop\GoogleDrive_0.1\GoogleDrive>cd C:\Users\jofontan\Desktop\GoogleDrive_0.1\GoogleDrive\

C:\Users\jofontan\Desktop\GoogleDrive_0.1\GoogleDrive>java -Dtalend.component.manager.m2.repository="C:\Users\jofontan\Desktop\GoogleDrive_0.1\GoogleDrive\..\lib" -Xms256M -Xmx1024M -cp classpath.jar; tp1_intro.googledrive_0.1.GoogleDrive --context=Default

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|                                                                                                     |
|                                                                                                     |
|-----+-----+-----+-----+-----+-----+-----+-----+-----+
|id                                     |name                                     |mimeType|modifiedTime   |size|kind   |trashed|parents                                     |webViewLink
|-----+-----+-----+-----+-----+-----+-----+-----+-----+
|1_yI_0q8CyZABvlix72FJISWCYMB9iONA|testFetch_csv.csv|text/csv|2020-02-10T00:55:19Z|190|drive#file|false|[1DtrUBaad4gWGr_lw328srGmGmbHvshA7]|https://drive.google.com/file/d/1_yI_0q8CyZABvlix72FJISWCYMB9iONA/view?usp=drivesdk
|1vBuIGat8mxQkLIOnhFpmnZT093qcR-tL|test_csv.csv|text/csv|2020-02-10T00:27:30Z|190|drive#file|false|[1DtrUBaad4gWGr_lw328srGmGmbHvshA7]|https://drive.google.com/file/d/1vBuIGat8mxQkLIOnhFpmnZT093qcR-tL/view?usp=drivesdk
|-----+-----+-----+-----+-----+-----+-----+-----+-----+

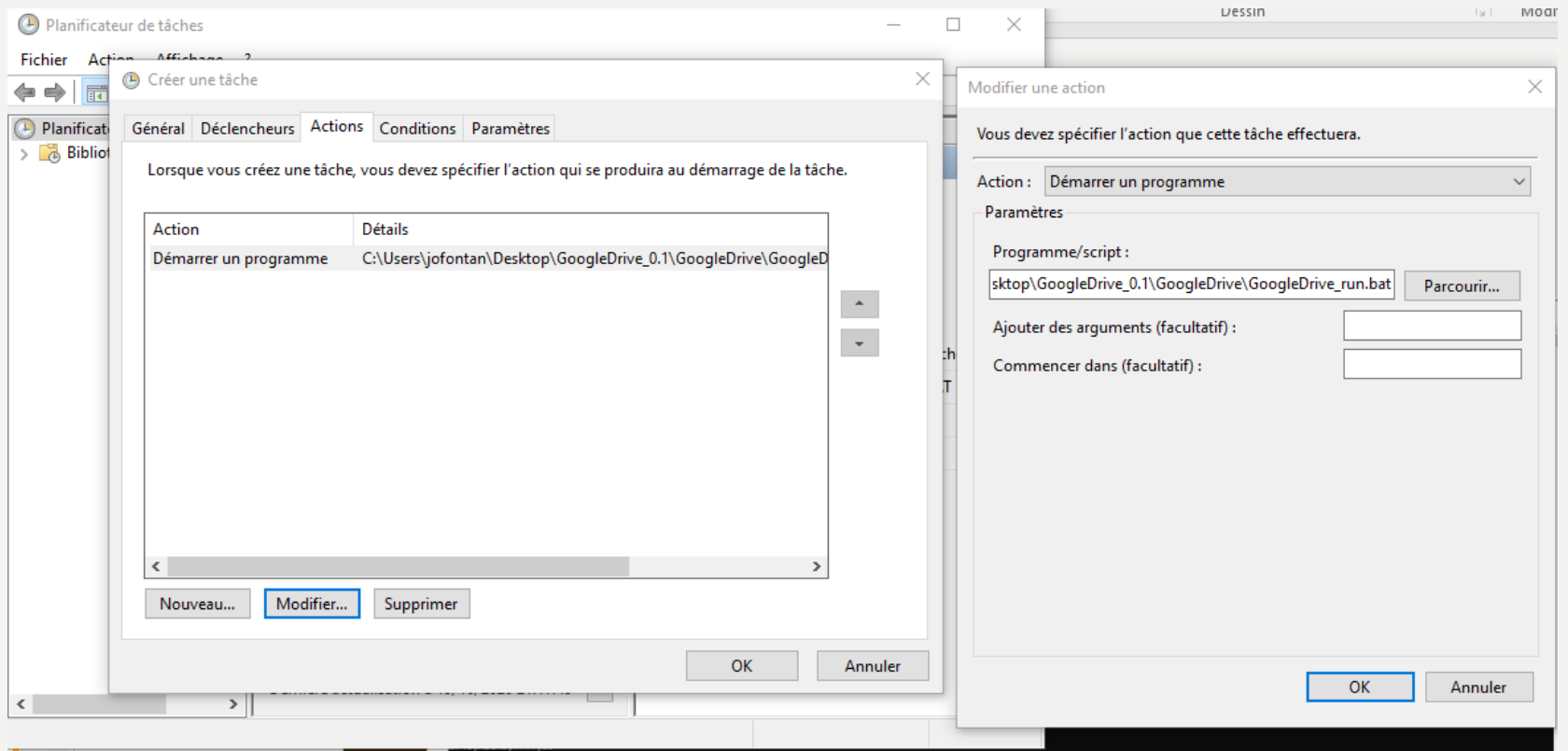
NUMERO;NOM;PRENOM;AGE
5484;DUPONT;TOTO;54
4654;DUPOND;TITI;36
84;TINTIN;TATA;49
8548;MILOU;TETE;25
484;TUNTUN;TUTU;41
8748;MILAU;TYTY;29
46854;TANTAN;TROTO;18
47874;TONTON;TRATA;52

NUMERO;NOM;PRENOM;AGE
5484;DUPONT;TOTO;54
4654;DUPOND;TITI;36
84;TINTIN;TATA;49
8548;MILOU;TETE;25
484;TUNTUN;TUTU;41
8748;MILAU;TYTY;29
46854;TANTAN;TROTO;18
47874;TONTON;TRATA;52
|1txhI6oCT3TnwaGeKFep1bM4-9Yr0uyZT|1HVbi-y35F8BSBPoleFBTsPkVSxliid18M

C:\Users\jofontan\Desktop\GoogleDrive_0.1\GoogleDrive>
```

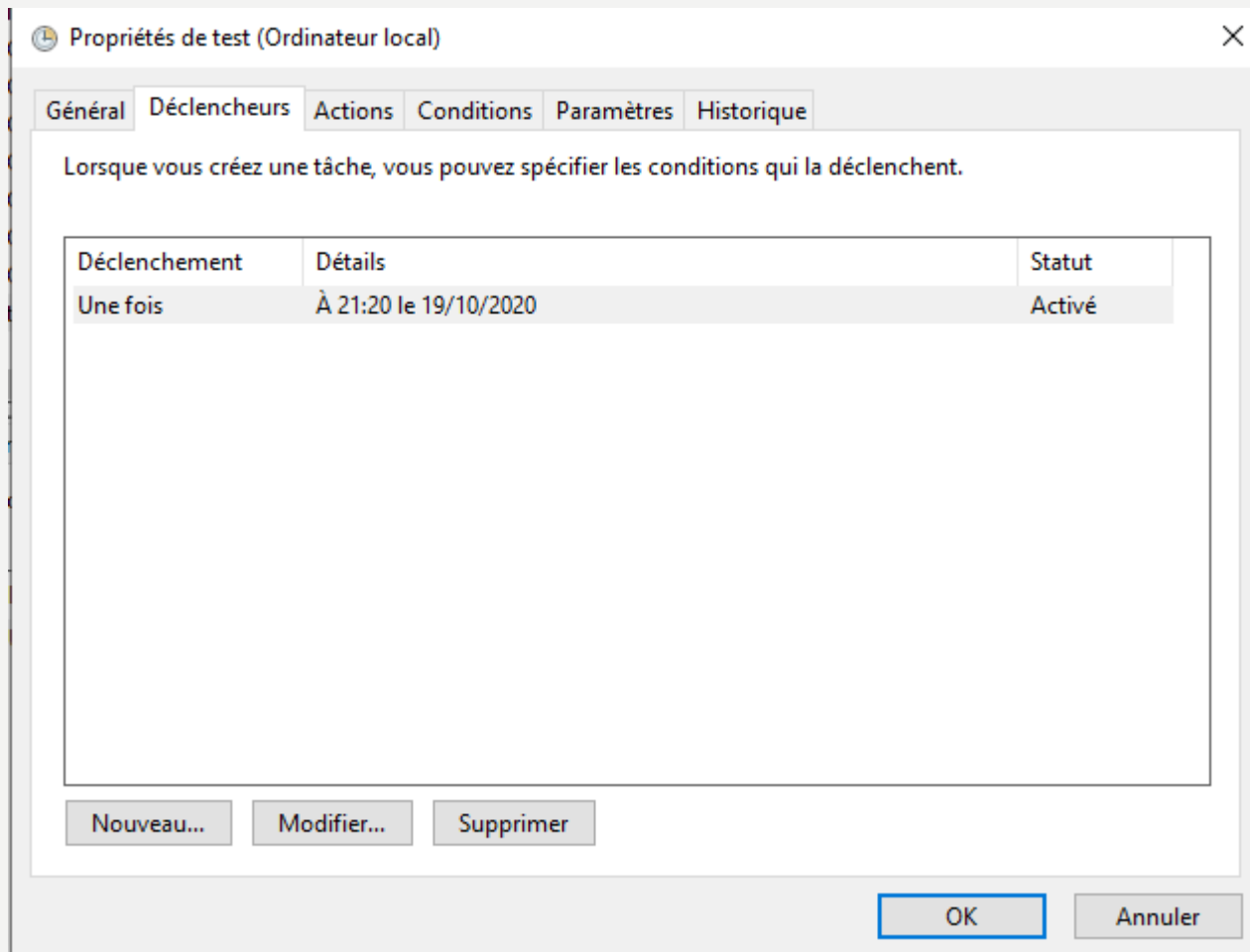
# ATELIERS SUR TALEND OPEN STUDIO

- Exporter le job pour programmer son lancement :
  - Créer via le planificateur de tâches, une tâche pour lancer votre script



# ATELIERS SUR TALEND OPEN STUDIO

- Exporter le job pour programmer son lancement :
  - Créer via le planificateur de tâches, une tâche pour lancer votre script



# ATELIERS SUR TALEND OPEN STUDIO

- Exporter le job pour programmer son lancement :
  - On peut voir l'exécution de la tâche



The screenshot displays the 'Planificateur de tâches' (Task Scheduler) window in Talend Open Studio. The interface includes a menu bar with 'Fichier', 'Action', and 'Affichage'. A left-hand pane shows a tree view of the task library, including folders for 'Microsoft', 'Mozilla', and 'MySQL'. The main area is titled 'Résumé du Planificateur de tâches (dernière réactualisation : 19/10/2020 21:20:19)'. It contains a section 'Vue d'ensemble du Planificateur de tâches' with a clock icon and text explaining that tasks are stored in the library and can be executed via the 'Actions' menu. Below this is a 'Statut de la tâche' section showing the status of tasks that started during a specific period: 'État des tâches qui ont démarré au cours de la période de temps suivante : Résumé : 75 au total - 1 exécution en cours, 66 réussite, 0 à l'arrêt, 8 échec'. At the bottom, a table lists individual tasks with their execution status and timing.

Nom de la tâche	Résultat d...	Démarrage de l'exécution	Fin de l'exécution	Déclenché par
TempSignedLicenseExchange (dernière exéc...				
test (dernière exécution réussi à 19/10/2020 ...)				
test	Réussite	19/10/2020 21:20:00	19/10/2020 21:20:05	Calendrier
UserTask (dernière exécution réussi à 19/10/...				
UserTask-Boom (dernière exécution réussi à ...)				

# ATELIERS SUR TALEND OPEN STUDIO

- Gestion des configurations via les contextes :
  - Reprendre le premier job puis éditer la connexion vers la base de données dans les métadonnées
  - On va ajouter des champs à gérer en fonction de la configuration. Pour cela, on va ouvrir la connexion vers la base de données puis cliquer sur « exporter en tant que contexte »

**Mettre à jour la connexion à une base de données - Etape 2/2**

 Vous devez cliquer sur le bouton Vérifier afin de vérifier les paramètres de la base de données. 

Type de BdD

Version de la base de données

Chaîne de caractères de connexion

Identifiant


Mot de passe

Serveur

Port

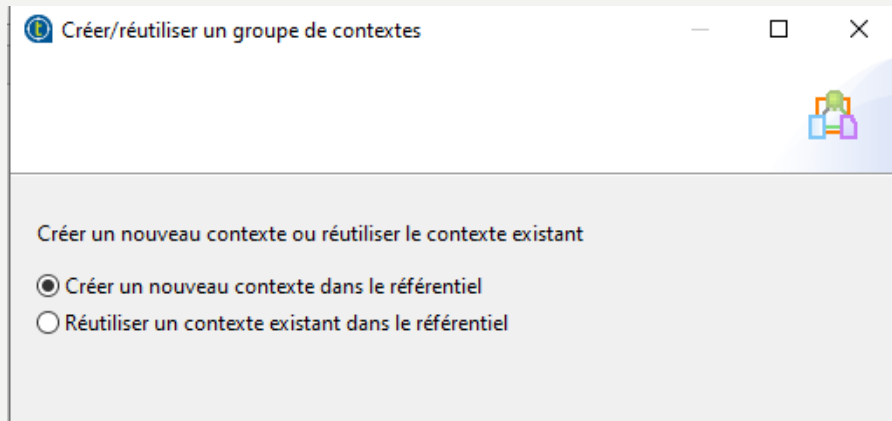
Base de données

Paramètres supplémentaires



# ATELIERS SUR TALEND OPEN STUDIO

- Créer deux configurations via les contextes :
  - Cliquer sur « créer un nouveau contexte dans le référentiel »



- Cliquer sur Next :

Créer/réutiliser un groupe de contextes

**Etape 1 sur 2**

Toute information requise

Nom: ConnectionMySQL

Objectif:

Description:

Créé par: user@talend.com

Verrouillé par:

Version: 0.1 [M] [m]

Statut:

Chemin d'accès: Sélectionner

< Back Next > Finish Cancel

# ATELIERS SUR TALEND OPEN STUDIO

- Créer deux configurations via les contextes :
  - On peut voir les paramètres qui sont possibles d'ajouter à la configuration.
  - On va ajouter deux contextes supplémentaires (2 configurations) avec le bouton + à droite de l'écran

Créer/réutiliser un groupe de contextes

Etape 2 sur 2

Définir les contextes, les variables et les valeurs

	Name	Type	Comment	Default		
				Value		
1						
2	ConnectionMySQL_Logi	String		access_bdd	<input type="checkbox"/>	
3	ConnectionMySQL_Pass	Password		*****	<input type="checkbox"/>	
4	ConnectionMySQL_Data	String		gestion_livres	<input type="checkbox"/>	
5	ConnectionMySQL_Serv	String		localhost	<input type="checkbox"/>	
6	ConnectionMySQL_Port	String		3307	<input type="checkbox"/>	
7	ConnectionMySQL_Addi	String		noDatetimeStringSync=true&server	<input type="checkbox"/>	

+ x

Default context environment Default

< Back Next > Finish Cancel

# ATELIERS SUR TALEND OPEN STUDIO

- Créer deux configurations via les contextes :
  - Créer deux nouveaux contextes puis appuyer sur OK :
    - PREPROD
    - PROD
  - Renommer le contexte Default par DEV

The image shows two overlapping windows from Talend Open Studio. The background window is titled 'Créer/réutiliser un groupe de contextes' and is on 'Etape 2 sur 2'. It contains a table with 7 rows of context variables. The foreground window is titled 'Configure Contexts' and shows a list of contexts: 'Default (Default)', 'PREPROD', and 'PROD'. The 'Default (Default)' context is selected with a checked checkbox. Buttons for 'Nouveau...', 'Editer...', 'Remove', 'Up', 'Descendre', 'OK', and 'Cancel' are visible.

	Name	Type	Comment	Default Value	
1					
2	ConnectionMySQL_Logi	String		access_bdd	<input type="checkbox"/>
3	ConnectionMySQL_Pass	Password		*****	<input type="checkbox"/>
4	ConnectionMySQL_Data	String		gestion_livres	<input type="checkbox"/>
5	ConnectionMySQL_Serv	String		localhost	<input type="checkbox"/>
6	ConnectionMySQL_Port	String		3307	<input type="checkbox"/>
7	ConnectionMySQL_Addi	String		noDatetimeStringSync=true&server	<input type="checkbox"/>



# ATELIERS SUR TALEND OPEN STUDIO

- Créer deux configurations via les contextes :
  - Vous obtenez les trois contextes suivants :

	Name	Type	Comment	DEV	PREPROD	PROD
				Value	Value	Value
1						
2	ConnectionMySQL_Logi	String		access_bdd	access_bdd	access_bdd
3	ConnectionMySQL_Pass	Password		*****	*****	*****
4	ConnectionMySQL_Data	String		gestion_livres	gestion_livres	gestion_livres
5	ConnectionMySQL_Serv	String		localhost	localhost	localhost
6	ConnectionMySQL_Port	String		3307	3307	3307
7	ConnectionMySQL_Addi	String		noDatetimeStringSync=true&server	noDatetimeStringSync=true&server	noDatetimeStringSync=true&server

- Votre connexion a changé, il faut faire la MAJ des jobs :

Connexion à la base de données

Mettre à jour la connexion à une base de données - Etape 2/2

Type de Bdd: MySQL

Version de la base de données: MySQL 8

Chaîne de caractères de connexion: jdbc:mysql://localhost:3307/gestion\_livres?noDatetimeStringSync=true&serverTimezone=UTC

Identifiant: context.ConnectionMySQL\_Login

Mot de passe: context.ConnectionMySQL\_Password

Serveur: context.ConnectionMySQL\_Server

Port: context.ConnectionMySQL\_Port

Base de données: context.ConnectionMySQL\_Database

Paramètres supplémentaires: context.ConnectionMySQL\_AdditionalParams

Ajouter un groupe de contextes

Ajouter le groupe de contexte inexistant du Job depuis le contexte du référentiel.

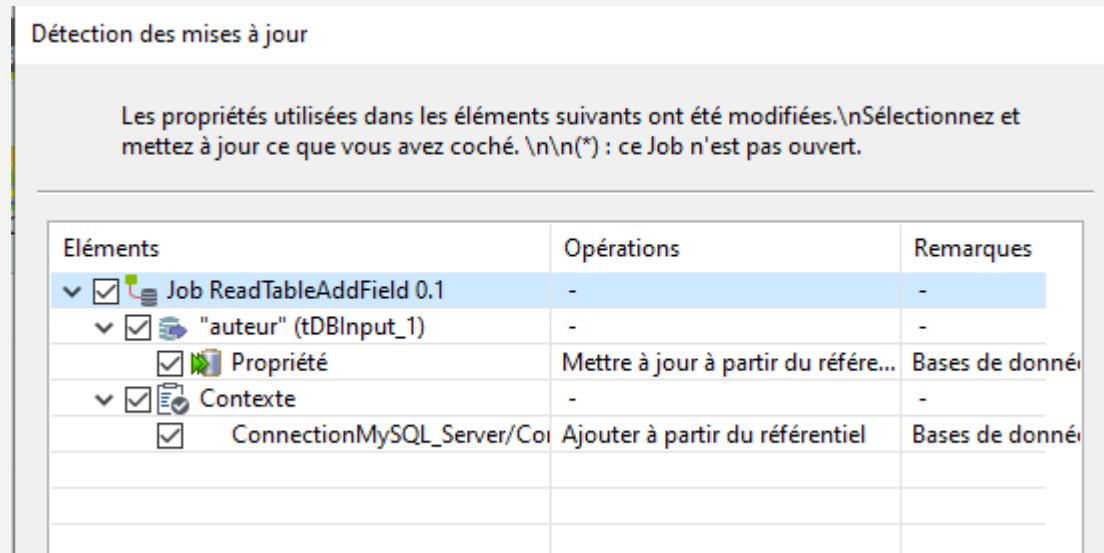
- Context: ConnectionMySQL
  - DEV (Défaut)

Sélectionner tout

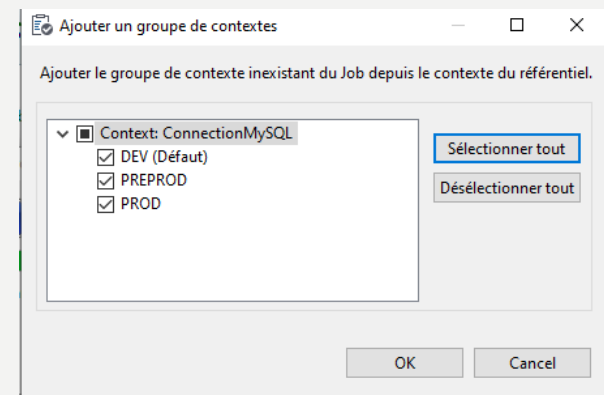
Désélectionner tout

# ATELIERS SUR TALEND OPEN STUDIO

- Créer deux configurations via les contextes :
  - Vous pouvez ouvrir le job de lecture d'une table, cliquer sur ok :



- Cliquer sur les trois environnements :



# ATELIERS SUR TALEND OPEN STUDIO

- Créer deux configurations via les contextes :
  - Modifier le contexte PREPROD pour aller chercher les données sur le PC du binôme à côté de vous. Pour cela, ouvrir le contexte « ConnectionMYSQL » :

Etape 2 sur 2  
Définir les contextes, les variables et les valeurs

	Name	Type	Comment	DEV	PREPROD	PROD		
				Value	Value	Value	<input type="checkbox"/>	<input type="checkbox"/>
1	ConnectionMySQL_Login	String		access_bdd	isima	access_bdd	<input type="checkbox"/>	<input type="checkbox"/>
2	ConnectionMySQL_Password	Password		*****	*****	*****	<input type="checkbox"/>	<input type="checkbox"/>
3	ConnectionMySQL_Database	String		gestion_livres	gestion_livres	gestion_livres	<input type="checkbox"/>	<input type="checkbox"/>
4	ConnectionMySQL_Server	String		localhost	a210pc04.local.isima.fr	localhost	<input type="checkbox"/>	<input type="checkbox"/>
5	ConnectionMySQL_Port	String		3307	3307	3307	<input type="checkbox"/>	<input type="checkbox"/>
6	ConnectionMySQL_AdditionalParameters	String		noDatetimeStringSync=true&server	noDatetimeStringSync=true&server	noDatetimeStringSync=true&server	<input type="checkbox"/>	<input type="checkbox"/>

Default context environment: DEV

< Back   Next >   Finish   Cancel

- Sur cet exemple, le contexte PREPROD a été modifié avec les bonnes informations de la base de données du PC04

# ATELIERS SUR TALEND OPEN STUDIO

- Créer deux configurations via les contextes :
  - Vous pouvez aller dans le menu « Executer » et voir les changements quand vous exécutez avec le contexte DEV et PREPROD

The screenshot displays the Talend Open Studio interface during a job execution. At the top, a job flow diagram shows three components: 'auteur', 'tMap\_1', and 'tLogRow\_1'. Performance metrics indicate '7 rows in 0,23s' and '29,91 rows/s' for both 'row1 (Main)' and 'sortie (Main)'. Below the diagram, the 'Designer' tab is active, showing the job configuration for 'Job ReadTableAddField 0.1'. The 'Contexts' dropdown is set to 'PREPROD'. The 'Exécution simple' section contains buttons for 'Exécuter', 'Arrêter', and 'Effacer'. A data table is visible, showing columns for 'id', 'id\_auteur', 'nom\_auteur', 'prenom\_auteur', 'ville', and 'date\_naissance'. To the right, the 'Contexts' table is displayed, showing various context variables and their values.

id	id_auteur	nom_auteur	prenom_auteur	ville	date_naissance
1	8547	Chambord	Emilie	Nice	Auteur ancien
0	52136	Fabiere	Sylvie	Bordeaux	Auteur ancien
0	78545	TINTIN	Thierry	Clermont	Auteur ancien
1	85478	Castafiore	Emilie	Paris	Auteur ancien
1	542536	Dupont	Pierre	Avignon	Auteur récent
0	6666666	BAMBI	NO	CF	Auteur récent
2	8547585	Momo	Roland	Toulouse	Auteur récent

Nom	Valeur
ConnectionMySQL...	isima
ConnectionMySQL...	****
ConnectionMySQL...	gestion_livres
ConnectionMySQL...	a210pc04.local.isi...
ConnectionMySQL...	3307
ConnectionMySQL...	noDatetimeString...

- A votre avis, à quoi peut servir cette fonctionnalité ?
- Exporter votre job avec le contexte DEV puis exécuter le à l'extérieur de Talend
- Exporter votre job avec le contexte PREPROD

# ATELIERS SUR TALEND OPEN STUDIO

- Créer deux configurations via les contextes :
  - Exporter votre job avec le contexte PREPROD

Construire le Job

Vers le fichier archive : C:\Talend\TOS\_DI-Win32-20200219\_1130-V7.3.1\ReadTableAddField\_0.1.zip

Job Version  
Sélectionnez la version de Job 0.1

Type de construction  
Sélectionner le type de construction Job standalone  Extraire le fichier zip

Options  
 Interpréteur de commande Tout  
 Scripts de contexte PREPROD  Appliquer le contexte aux Jobs fils

Eléments  
 Sources Java

# ATELIERS SUR TALEND OPEN STUDIO

- Créer deux configurations via les contextes :
  - Executer votre job avec le contexte PREPROD

```
c:\Talend\ReadTableAddField_0.1\ReadTableAddField>ReadTableAddField_run.bat
c:\Talend\ReadTableAddField_0.1\ReadTableAddField>c:
c:\Talend\ReadTableAddField_0.1\ReadTableAddField>cd c:\Talend\ReadTableAddField_0.1\ReadTableAddField\
c:\Talend\ReadTableAddField_0.1\ReadTableAddField>java -Dtalend.component.manager.m2.repository="c:\Talend\ReadTableAddField_0.1\ReadTableAddField\..\lib" -Xms256M -Xmx1024M -cp .;..\lib\routines.jar;..\lib\log4j-to-slf4j-2.12.1.jar;..\lib\slf4j-log4j12-1.7.25.jar;..\lib\log4j-1.2.17.jar;..\lib\mysql-connector-java-8.0.18.jar;..\lib\crypto-utils.jar;..\lib\slf4j-api-1.7.25.jar;..\lib\dom4j-2.1.1.jar;readtableaddfield_0_1.jar; tp1.readtableaddfield_0_1.ReadTableAddField --context=PREPROD
```

tLogRow_1						
Yacouba	NUMERO_A	NOM	PRENOM	DOMICILE	var1	
1	8547	Chambord	Emilie	Nice	Auteur	ancien
0	52136	Fabiere	Sylvie	Bordeaux	Auteur	ancien
0	78545	TINTIN	Thierry	Clermont	Auteur	ancien
1	85478	Castafiore	Emilie	Paris	Auteur	ancien
1	542536	Dupont	Pierre	Avignon	Auteur	rúcent
0	6666666	BAMBI	NO	CF	Auteur	rúcent
2	8547585	Momo	Roland	Toulouse	Auteur	rúcent

- Changer le fichier .bat pour faire appel au contexte DEV :

# ATELIERS SUR TALEND OPEN STUDIO

- Créer deux configurations via les contextes :
  - Editer le fichier de votre job exporté avec le contexte de PREPROD pour appliquer le contexte DEV

```
ReadTableAddField_run.bat
1  %~d0
2  cd %~dp0
3  java -Dtalend.component.manager.m2.repository="%cd%/../lib" -Xms256M -Xmx1024M -cp
   ./../lib/routines.jar;../lib/log4j-to-slf4j-2.12.1.jar;../lib/slf4j-log4j12-1.7.25.jar;../lib/log4j-1.2.17.jar;../lib/mysql-conn
   ector-java-8.0.18.jar;../lib/crypto-utils.jar;../lib/slf4j-api-1.7.25.jar;../lib/dom4j-2.1.1.jar;readtableaddfield_0_1.jar;
   tpl.readtableaddfield_0_1.ReadTableAddField --context=DEV %*
```

- Exécuter le job à nouveau :

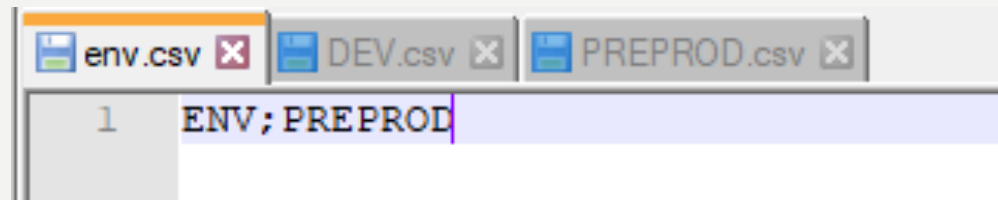
```
c:\Talend\ReadTableAddField_0.1\ReadTableAddField>ReadTableAddField_run.bat
c:\Talend\ReadTableAddField_0.1\ReadTableAddField>c:
c:\Talend\ReadTableAddField_0.1\ReadTableAddField>cd c:\Talend\ReadTableAddField_0.1\ReadTableAddField\
c:\Talend\ReadTableAddField_0.1\ReadTableAddField>java -Dtalend.component.manager.m2.repository="c:\Talend\ReadTableAddF
ield_0.1\ReadTableAddField/../lib" -Xms256M -Xmx1024M -cp ./../lib/routines.jar;../lib/log4j-to-slf4j-2.12.1.jar;../lib/
slf4j-log4j12-1.7.25.jar;../lib/log4j-1.2.17.jar;../lib/mysql-connector-java-8.0.18.jar;../lib/crypto-utils.jar;../lib/s
lf4j-api-1.7.25.jar;../lib/dom4j-2.1.1.jar;readtableaddfield_0_1.jar; tp1.readtableaddfield_0_1.ReadTableAddField --con
text=DEV

-----+-----+-----+-----+-----+-----+-----+
                          tLogRow_1
-----+-----+-----+-----+-----+-----+-----+
|Yacouba|NUMERO_A|NOM      |PRENOM |DOMICILE|var1  |
-----+-----+-----+-----+-----+-----+-----+
|1      |8547   |Chambord |Emilie |Nice    |Auteur ancien|
|0      |52136  |Fabiere  |Sylvie |Bordeaux|Auteur ancien|
|0      |78545  |TINTIN   |Thierry|Clermont|Auteur ancien|
|1      |85478  |Castafione|Emilie |Paris   |Auteur ancien|
|1      |542536 |Dupont   |Pierre |Avignon |Auteur rúcent|
|2      |8547585|Momo     |Roland |Toulouse|Auteur rúcent|
-----+-----+-----+-----+-----+-----+-----+

```

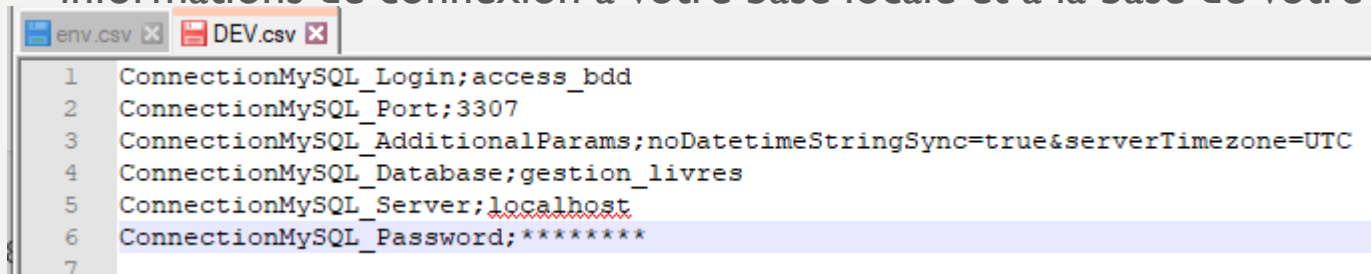
# ATELIERS SUR TALEND OPEN STUDIO

- Exécution du job standalone en fonction de l'environnement :
  - Créer un fichier csv qui se nomme env.csv, il contiendra les environnements.
  - Pour changer, il suffira d'écrire DEV à la place de PREPROD

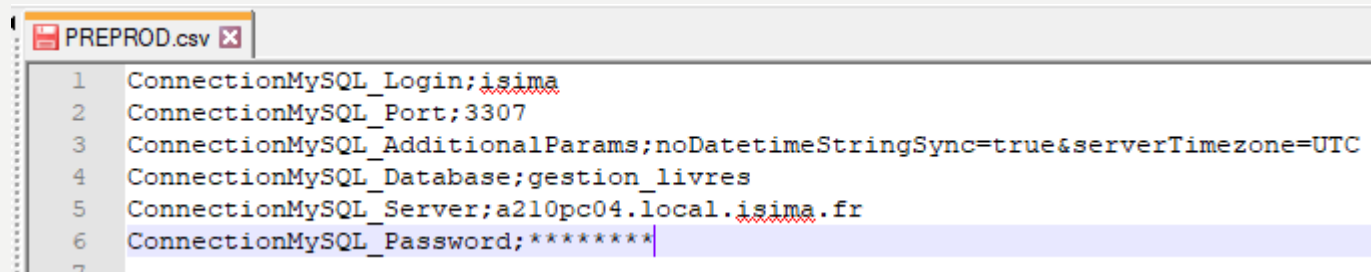


```
env.csv x DEV.csv x PREPROD.csv x
1 ENV; PREPROD
```

- Créer deux fichiers de contexte DEV.csv et PREPROD.csv reprenant les informations de connexion à votre base locale et à la base de votre binôme :



```
env.csv x DEV.csv x
1 ConnectionMySQL_Login;access_bdd
2 ConnectionMySQL_Port;3307
3 ConnectionMySQL_AdditionalParams;noDatetimeStringSync=true&serverTimezone=UTC
4 ConnectionMySQL_Database;gestion_livres
5 ConnectionMySQL_Server;localhost
6 ConnectionMySQL_Password;*****
7
```

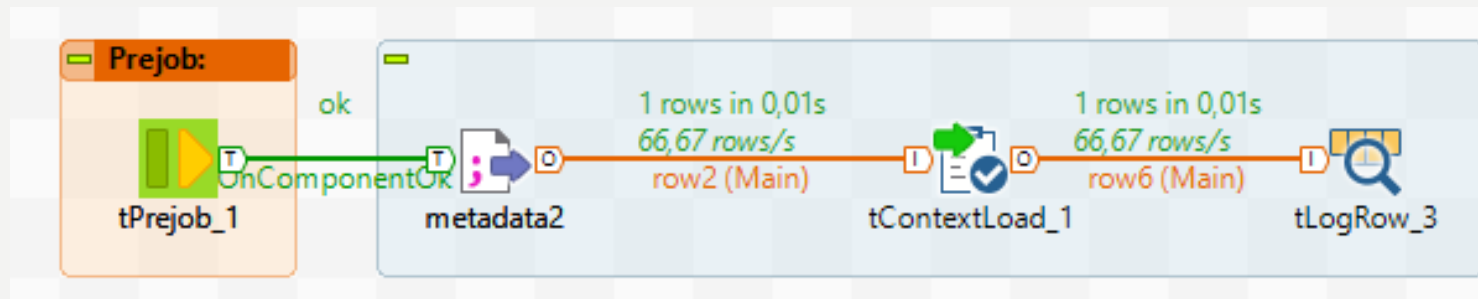


```
PREPROD.csv x
1 ConnectionMySQL_Login;isima
2 ConnectionMySQL_Port;3307
3 ConnectionMySQL_AdditionalParams;noDatetimeStringSync=true&serverTimezone=UTC
4 ConnectionMySQL_Database;gestion_livres
5 ConnectionMySQL_Server;a210pc04.local.isima.fr
6 ConnectionMySQL_Password;*****
7
```

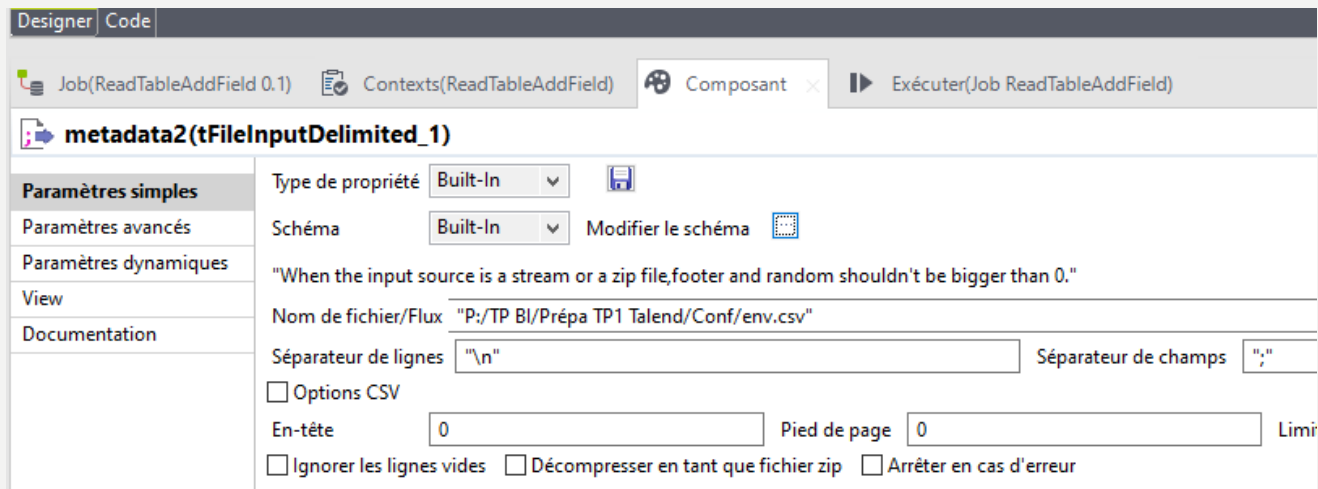


# ATELIERS SUR TALEND OPEN STUDIO

- Exécution du job standalone en fonction de l'environnement :
  - Créer un préjob dans votre job de lecture de base de données

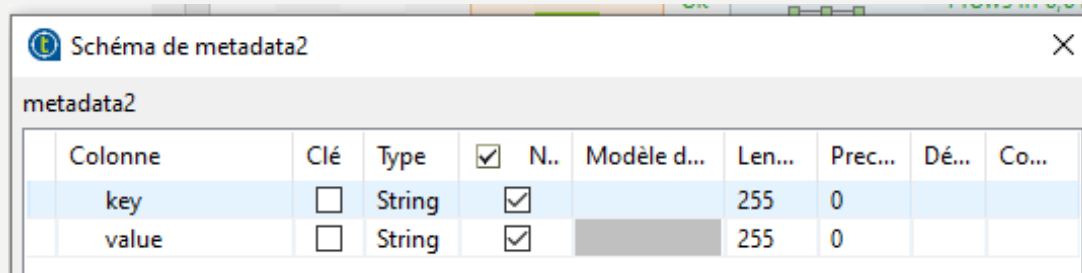


- Pour cela, il faut faire glisser un composant tPrejob dans le designer puis relier un tFileInputDelimited dans l'interface en remplaçant le nom de fichier par l'endroit où vous avez stocké votre fichier :



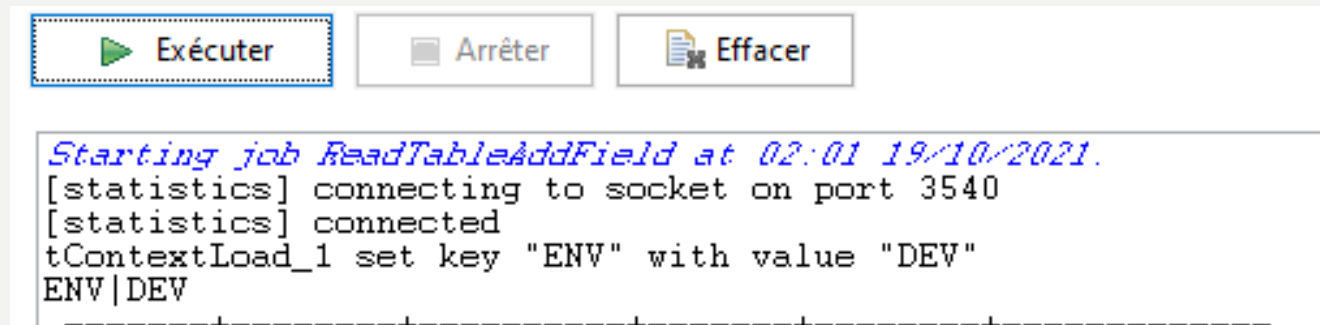
# ATELIERS SUR TALEND OPEN STUDIO

- Exécution du job standalone en fonction de l'environnement :
  - Remplir pour ce tFileInputDelimited, le schéma suivant :



Colonne	Clé	Type	<input checked="" type="checkbox"/> N..	Modèle d...	Len...	Prec...	Dé...	Co...
key	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		255	0		
value	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		255	0		

- Relier le composant tContextLoad en cochant « Afficher les opérations »
- Relier un composant tlogRow
- Exécuter votre job

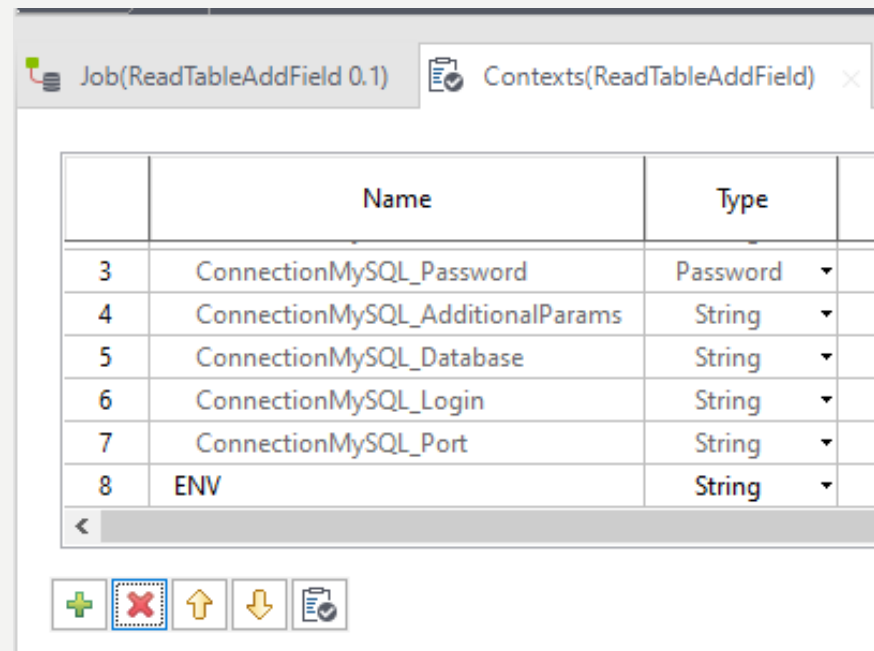


Exécuter Arrêter Effacer

```
Starting job ReadTableAddField at 02:01 19/10/2021.  
[statistics] connecting to socket on port 3540  
[statistics] connected  
tContextLoad_1 set key "ENV" with value "DEV"  
ENV|DEV
```

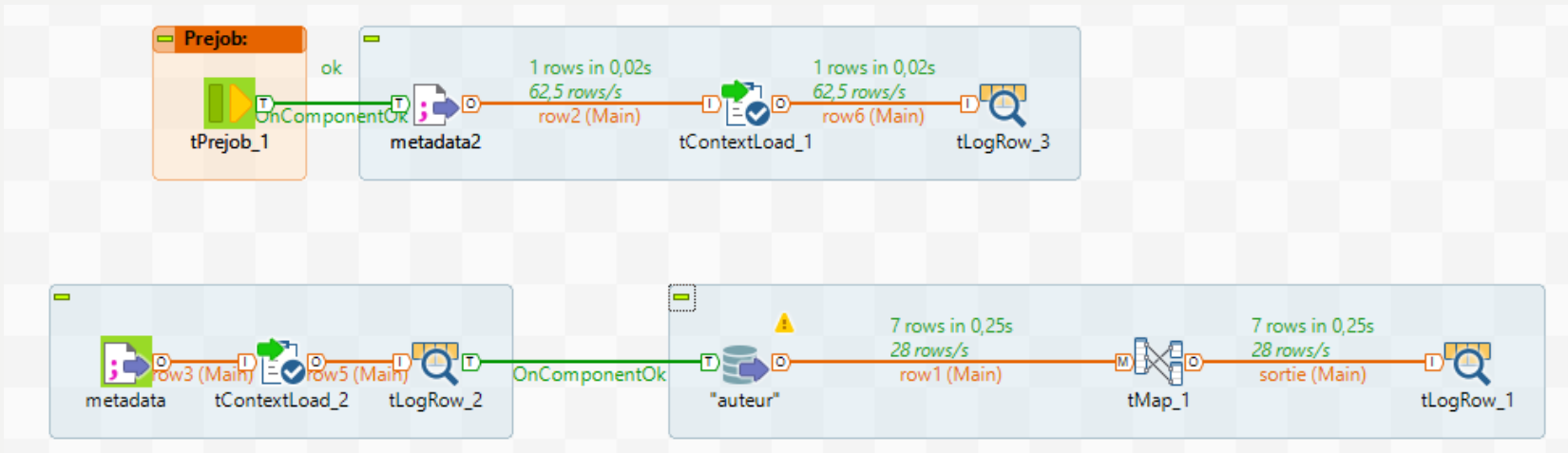
# ATELIERS SUR TALEND OPEN STUDIO

- Exécution du job standalone en fonction de l'environnement :
  - Créer la variable locale de contexte : ENV.
  - Pour cela, il faut se rendre dans « Contexts » à côté « Exécuter » en appuyant sur la flèche « + »



# ATELIERS SUR TALEND OPEN STUDIO

- Exécution du job standalone en fonction de l'environnement :
  - Lecture du fichier de contexte en fonction de la valeur contenue dans le fichier ENV.csv



- Ajouter un composant tFileInputDelimited

# ATELIERS SUR TALEND OPEN STUDIO

- Exécution du job standalone en fonction de l'environnement :
  - Ajouter un composant tFileInputDelimited dans l'interface en remplaçant le nom de fichier par ce qui est écrit dans la capture (attention au début du chemin, cela dépend de l'endroit où vous avez stocké votre fichier de contexte)

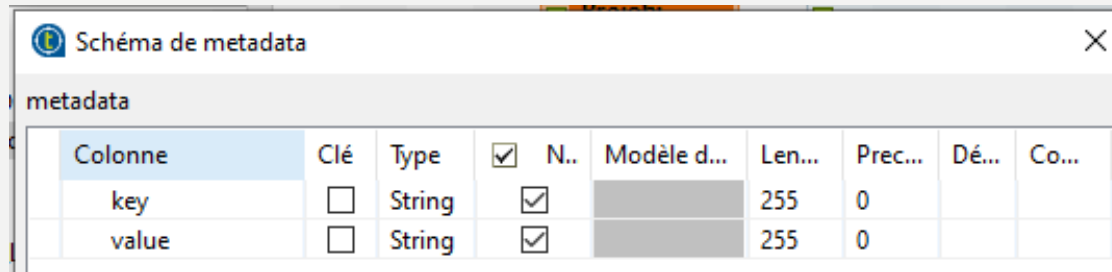
The screenshot displays the Talend Open Studio interface during a job execution. The top part shows a job graph with components: metadata, tContextLoad\_2, tLogRow\_2, "auteur", and tMap\_1. The "auteur" component is highlighted with a yellow warning icon and displays execution statistics: "7 rows in 0,25s" and "28 rows/s". The job is running in "row1 (Main)".

The bottom part shows the configuration window for the "metadata(tFileInputDelimited\_2)" component. The configuration is as follows:

Section	Configuration
Type de propriété	Built-In
Schéma	Built-In
Message d'erreur	"When the input source is a stream or a zip file, footer and random shouldn't be bigger than 0."
Nom de fichier/Flux	"P:/TP BI/Prépa TP1 Talend/Conf/" + context.ENV + ".csv"
Séparateur de lignes	"\n"
Séparateur de champs	","
Options CSV	<input type="checkbox"/>
En-tête	0
Pied de page	0
Ignorer les lignes vides	<input type="checkbox"/>
Décompresser en tant que fichier zip	<input type="checkbox"/>
Arrêter en cas d'erreur	<input type="checkbox"/>

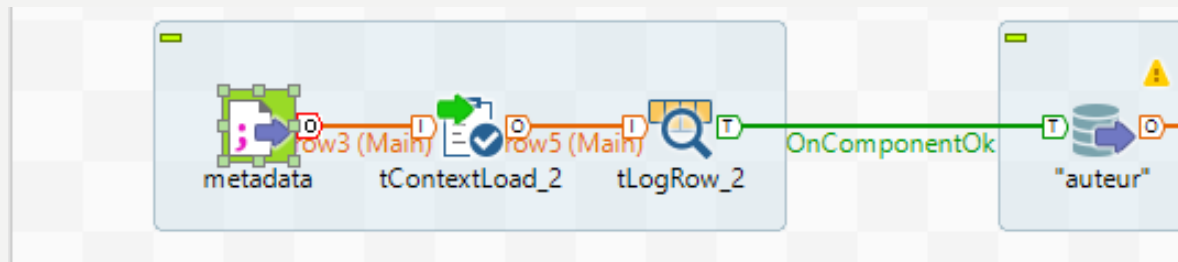
# ATELIERS SUR TALEND OPEN STUDIO

- Exécution du job standalone en fonction de l'environnement :
  - Modifier le schéma pour que celui-ci soit identique à la capture :



Colonne	Clé	Type	<input checked="" type="checkbox"/> N..	Modèle d...	Len...	Prec...	Dé...	Co...
key	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		255	0		
value	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		255	0		

- Relier le précédent composant à un composant tContextLoad en cochant « Afficher les opérations »
- Relier le précédent composant à un composant tLogRow
- Relier le précédent composant à la lecture de base :



# ATELIERS SUR TALEND OPEN STUDIO

- Exécution du job standalone en fonction de l'environnement :
  - Exécuter le job avec la valeur DEV dans le fichier env.csv
  - Exécuter le job avec la valeur PREPROD dans le fichier env.csv
  - Que se passe-t-il ?
  - Essayer de modifier le contexte sur la partie droite d'exécuter
  - Que se passe-t-il ?

The screenshot shows the Talend Open Studio interface during a job execution. The main window displays a data table with the following content:

1	8547	Chambord	Emilie	Nice	Auteur ancien	
0	52136	Fabiere	Sylvie	Bordeaux	Auteur ancien	
0	78545	TINTIN	Thierry	Clermont	Auteur ancien	
1	85478	Castafiore	Emilie	Paris	Auteur ancien	
1	542536	Dupont	Pierre	Avignon	Auteur récent	
0	6666666	BAMBI	NO	CF	Auteur récent	
2	8547585	Momo	Roland	Toulouse	Auteur récent	

Below the table, the text "[statistics] disconnected" is visible. To the right, a context configuration window is open, showing the "PREPROD" environment. It contains a table with the following data:

Nom	Valeur
ConnectionMySQL...	isima
ConnectionMySQL...	****
ConnectionMySQL...	gestion_livres
ConnectionMySQL...	a210pc04.local.isi...
ConnectionMySQL...	3307
ConnectionMySQL...	noDatetimeString...
ENV	

# ATELIERS SUR TALEND OPEN STUDIO

- Exécution du job standalone en fonction de l'environnement :
  - Exporter le job comme précédent via la construction de job
  - Exécution avec le fichier env.csv paramétré avec DEV
  - Exécution avec le fichier env.csv paramétré avec PREPROD
  - Imaginez des cas d'utilisation pour ce job ?



# ATELIERS SUR TALEND OPEN STUDIO

- Exécution du job standalone en fonction de l'environnement :
  - Exercice : Associer l'environnement de PROD à une des bases Clevercloud.com
  - Exporter votre job
  - Ordonnancer votre job avec le contexte PROD

NUMERO_A	NOM	PRENOM	DOMICILE	var1
1	Taylor	Annie	PenArgyl	Auteur ancien
2	Marine	Corbin	Nice	Auteur ancien
3	Jeanette	Charron	Agen	Auteur ancien
4	Antoine	Phaneuf	Epinal	Auteur ancien
5	Philippine	Rochefort	SAINT-LAURENT-DU-VAR	Auteur ancien
6	Florus	Racine	NOGENT-SUR-MARNE	Auteur ancien