

Mini-rapport – Chaîne Talend “Paie → Contrôle → Journaux & KPI”

1) Contexte & objectif

Mettre en place un flux Talend qui :

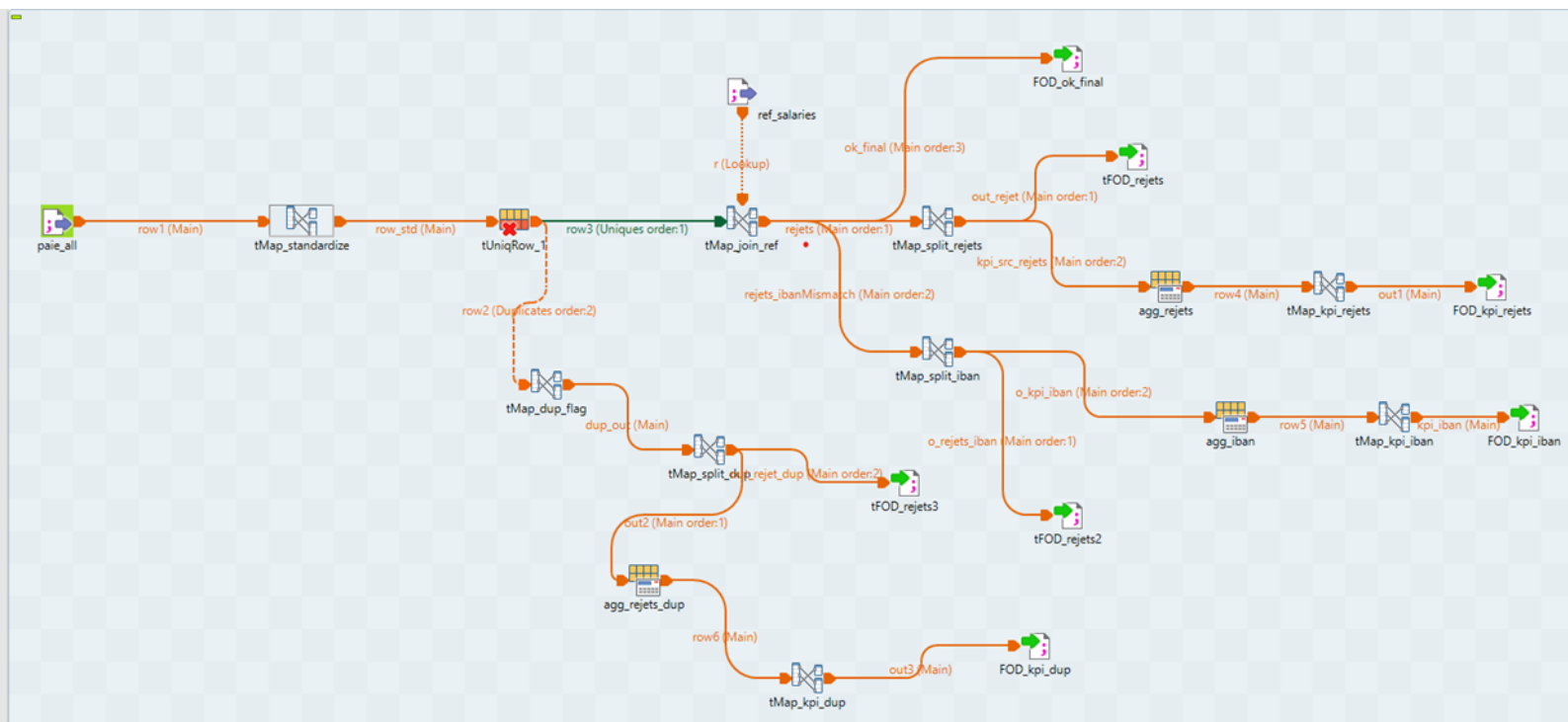
- **ingère la paie mensuelle,**
- **standardise et dédoublonne** les lignes,
- **croise** avec un **référentiel salariés,**
- applique des **règles de qualité** (formats, montants, IBAN...),
- **sépare** les **lignes valides** des **rejets**,
- produit des **journaux de rejets** et des **KPI d'exécution**.

Résultat : des **fichiers plats** destinés aux applications aval (compta, contrôle interne, pilotage).

2) Données d'entrée

- **rh_paie**: flux principal de paie (27 lignes dans l'exemple).
- **ref_salaries** : référentiel (lookup) sur **matricule** (nom, prénom, IBAN, centre de coût...).

3) Vue d'ensemble du job (capture)



Chaîne principale de gauche à droite :

1. **tFileInputDelimited paie_all**
Lecture CSV des paies.
2. **tMap tMap_standardize**
Mise au propre des types et formats (dates, numériques, trimming).
3. **tUniqRow tUniqRow_1**
Dédoublonnage sur une clé logique (mois + matricule).
 - **Main (uniques)** → continue le traitement.
 - **Reject/Duplicates** → sortie "row2".
4. **tMap tMap_dup_flag** (voie doublons)
Marquage et formatting des doublons → **dup_out**
 - **tMap_split_dup_rejet_dup** :
 - **Main (order:2)** → **tFOD_rejets3** (journal doublons)
 - **out2 (order:1)** → **agg_rejets_dup** → **tMap_kpi_dup** → **FOD_kpi_dup**
5. **tMap tMap_join_ref** (voie uniques)
Join avec ref_salaries (clé : matricule).
Évaluation des règles (ex. format IBAN, salaires/heures/présence en référentiel).
3 sorties ordonnées :
 - **ok_final (Main order:3)** → **FOD_ok_final** (lignes prêtes pour la suite)
 - **rejets (Main order:1)** → **tMap_split_rejets**
 - **kpi_src_rejets** → **agg_rejets** → **tMap_kpi_rejets** → **FOD_kpi_rejets**
 - en parallèle **tFOD_rejets** (journal rejets "règles")
 - **rejets_ibanMismatch (Main order:2)** → **tMap_split_iban**
 - **o_rejets_iban (Main order:1)** → **agg_iban** → **tMap_kpi_iban** → **FOD_kpi_iban**
 - en parallèle **o_kpi_iban (order:2)** → **tFOD_rejets2** (journal IBAN)

4) Règles de qualité appliquées (exemples)

- **Dédoublonnage** : même mois + même matricule.
- **Référentiel** : matricule connu dans ref_salaries.
- **Formats** : IBAN conforme (regex pays), valeurs numériques ≥ 0.
- **Cohérence** : salaires ou heures non négatifs, etc.

Les lignes qui violent une règle sont routées en rejets avec un **rule_id** et un **rule_msg** (ex. Q07_IBAN_MISMATCH, Q05_DUPLICATE, Q00_INVALIDE).

5) Sorties produites

Journaux (CSV)

- **FOD_rejets** : rejets "règles" (ex. montants invalides, absence référentiel...)
- **FOD_rejets2** : rejets **IBAN mismatch** (différence paie vs référentiel)
- **FOD_rejets3** : **doublons**
- **FOD_ok_final** : **lignes conformes** (prêtes à export vers la compta)

KPI d'exécution (CSV)

- **FOD_kpi_rejets** : REJETS_REGLES, <nb>, <horodatage>
 - **FOD_kpi_iban** : REJETS_IBAN_MISMATCH, <nb>, <horodatage>
 - **FOD_kpi_dup** : REJETS_DUPLONS, <nb>, <horodatage>
-

6) Paramétrage clé (bonnes pratiques)

- **Encodage** : UTF-8, séparateur ; (ou , selon le SI), **header** activé.
 - **tUniqRow** : cocher *Use selected column(s) for key* (clé mois+matricule).
 - **tMap (join)** :
 - Lookup ref_salaries en *Load Once, Unique match*.
 - Règles implémentées par expressions/colonnes techniques (isIbanFmtOK, etc.).
 - **Composants FOD** :
 - Activer **Create directory if not exists**.
 - Pour KPI : **Append** = true sur 2/3 composants, entête uniquement sur le premier.
 - **Robustesse** : tracer via tLogRow (mode Table) au pied de chaque branche si besoin.
-

7) Résultats (jeu d'essai de la capture)

- **Entrée** : 27 lignes.
 - **Uniques** : 22 lignes. **Doublons** : 5 lignes.
 - **Rejets "règles"** : ~1 ligne.
 - **Rejets IBAN mismatch** : ~1 ligne.
 - **OK final** : 18 lignes
 - **KPI** : 3 lignes ajoutées dans le fichier KPI (1 par famille de rejets).
-

9) Évolutions possibles

- **Paramétrage par contextes** (répertoires in/out, noms de fichiers datés).
- **Agrégation KPI unique** (un seul fichier partagé par les trois branches).
- **Chargement base de données** en sortie (audits, historisation).

ANNEXE

rh_paie.csv

	A	B	J	K	L	M	N	O	P	Q	R	S	T	U
1	periode	matricule	date_sortie	temps_travail_pct	salaire_base	primes	heures_sup	retenues	iban	bic	montant_heures_sup	brut	charges_patronales	net_a_payer
2	01/03/2025	E001			100 3200.0	250.0	120.0	180.0	FR7612345987650123456789014	AGRIFRPP	2400.0	5850.0	2457.0	5670.0
3	01/03/2025	E002			100 3700.0	0.0	60.0	220.0	FR1412345987650123456789012	BNPAFRPP	1200.0	4900.0	2058.0	4680.0
4	01/03/2025	E003			80 2400.0	100.0	0.0	120.0	FR3512345987650123456789015	SOGEFRPP	0.0	2500.0	1050.0	2380.0
5	01/03/2025	E004			100 4500.0	450.0	0.0	350.0	FR0212345987650123456789001	PSSTFRPP	0.0	4950.0	2079.0	4600.0
6	01/03/2025	E005			100 3000.0	0.0	30.0	150.0	FR4412345987650123456789019	CMCIFR2A	600.0	3600.0	1512.0	3450.0
7	01/03/2025	E006			100 3800.0	200.0	0.0	240.0	FR9112345987650123456789026	CCBPFRPP	0.0	4000.0	1680.0	3760.0
8	01/03/2025	E007			100 4200.0	350.0	0.0	260.0	FR1212345987650123456789003	CEPAFRPP	0.0	4550.0	1911.0	4290.0
9	01/03/2025	E008			100 4700.0	500.0	0.0	380.0	FR6912345987650123456789021	CMCIFRPP	0.0	5200.0	2184.0	4820.0
10	01/02/2025	E001			100 3200.0	200.0	60.0	170.0	FR7612345987650123456789014	AGRIFRPP	1200.0	4600.0	1932.0	4430.0
11	01/02/2025	E002			100 3700.0	0.0	0.0	210.0	FR1412345987650123456789012	BNPAFRPP	0.0	3700.0	1554.0	3490.0
12	01/02/2025	E003			80 2400.0	80.0	0.0	120.0	FR3512345987650123456789015	SOGEFRPP	0.0	2480.0	1041.6	2360.0
13	01/02/2025	E004			100 4500.0	300.0	0.0	340.0	FR0212345987650123456789001	PSSTFRPP	0.0	4800.0	2016.0	4460.0
14	01/04/2025	E002			100 3700.0	0.0	0.0	210.0	FR1412345987650123456789012	BNPAFRPP	0.0	3700.0	1554.0	3490.0
15	01/04/2025	E004			100 4500.0	300.0	0.0	340.0	FR0212345987650123456789001	PSSTFRPP	0.0	4800.0	2016.0	4460.0
16	01/04/2025				100 3200.0	250.0	120.0	180.0	FR7612345987650123456789014	AGRIFRPP	2400.0	5850.0	2457.0	5670.0
17	01/04/2025	E006			100 3800.0	200.0	0.0	240.0	FR9112345987650123456789026	CCBPFRPP	0.0	4000.0	1680.0	3760.0
18	01/04/2025	E006			100 3800.0	200.0	0.0	240.0	FR9112345987650123456789026	CCBPFRPP	0.0	4000.0	1680.0	3760.0
19	01/04/2025	E001			100 3200.0	200.0	60.0	170.0	FR7612345987650123456789014	AGRIFRPP	1200.0	4600.0	1932.0	4430.0
20	01/04/2025	E003			80 -2500.0	100.0	0.0	120.0	FR3512345987650123456789015	SOGEFRPP	0.0	-2400.0	-1008.0	-2520.0
21	01/04/2025	E002			100 3700.0	0.0	60.0	220.0		BNPAFRPP	1200.0	4900.0	2058.0	4680.0
22	01/04/2025	E999			100 3200.0	250.0	120.0	180.0	FR5512345987650123456789033	AGRIFRPP	2400.0	5850.0	2457.0	5670.0
23	01/04/2025	E005			100 3000.0	0.0	30.0	150.0	FR4412345987650123456789019	CMCIFR2A	600.0	3600.0	1512.0	3450.0
24	01/04/2025	E008			100 4700.0	500.0	0.0	380.0	FR6912345987650123456789021	CMCIFRPP	0.0	5200.0	2184.0	4820.0
25	01/04/2025	E003			80 2400.0	80.0	0.0	120.0	FR3512345987650123456789015	SOGEFRPP	0.0	2480.0	1041.6	2360.0
26	01/04/2025	E005			100 3000.0	0.0	30.0	150.0	FR4412345987650123456789019	CMCIFR2A	600.0	3600.0	1512.0	3450.0
27	01/04/2025	E004			100 4500.0	450.0	9999.0	350.0	FR0212345987650123456789001	PSSTFRPP	199980.0	204930.0	86070.6	204580.0
28	01/04/2025	E007			100 4200.0	350.0	0.0	260.0	FR00INVALIDIBAN0000000000000	CEPAFRPP	0.0	4550.0	1911.0	4290.0
29														
30														

ref_salaries.csv

	A	B	C	D	E	F
1	matricule	nom	prenom	iban	centre_co	site
2	E001	Martin	Alice	FR7612345987650123456789014	CC100	Paris
3	E002	Dupont	Benoit	FR1412345987650123456789012	CC200	Lyon
4	E003	Nguyen	Camille	FR3512345987650123456789015	CC300	Paris
5	E004	Lopez	Diego	FR0212345987650123456789001	CC100	Paris
6	E005	Bernard	Emma	FR4412345987650123456789019	CC400	Marseille
7	E006	Omar	Farid	FR9112345987650123456789026	CC200	Lille
8	E007	Garcia	Hugo	FR1212345987650123456789003	CC200	Lyon
9	E008	Ibrahim	Inès	FR6912345987650123456789021	CC100	Paris

	A	B	C
1	etape	nb_lignes	moment
2	REJETS_REGLES	1	03/08/2025 16:11
3	REJETS_IBAN_MISMATCH	1	03/08/2025 16:11
4	REJETS_DUPLONS	5	03/08/2025 16:11

	A	B	C	D	E	F	G	H	I
1	periode	matricule	centre_cout	salaire_base	heures_sup	iban	key_dedup	rule_id	rule_msg
2	01/04/2025	E003	CC300	-2500.0	0.0	FR3512345987650123456789015	2025-04_E003	Q03_SALAIRE_NEG	
3	01/04/2025	E007	CC200	4200.0	0.0	FR00INVALIDIBAN00000000000000	2025-04_E007	Q07_IBAN_MISMATCH	IBAN différent du référentiel
4	01/04/2025	E006	CC200	3800.0	3800.0	FR9112345987650123456789026	2025-04_E006	Q05_DUPLICATE	Doublon (2025-04+E006)
5	01/04/2025	E002	CC200	3700.0	3700.0	FR1412345987650123456789012	2025-04_E002	Q05_DUPLICATE	Doublon (2025-04+E002)
6	01/04/2025	E003	CC300	2400.0	2400.0	FR3512345987650123456789015	2025-04_E003	Q05_DUPLICATE	Doublon (2025-04+E003)
7	01/04/2025	E005	CC400	3000.0	3000.0	FR4412345987650123456789019	2025-04_E005	Q05_DUPLICATE	Doublon (2025-04+E005)
8	01/04/2025	E004	CC100	4500.0	4500.0	FR0212345987650123456789001	2025-04_E004	Q05_DUPLICATE	Doublon (2025-04+E004)

[illegible]