



# Extract, Transform, Load (ETL)

Automatisation, traçabilité, règles métiers, intégration

# Moteur d'intégration des flux

- DSIPilot Version 8.5 est un ETL développé en JAVA et une JVM OpenJDK 17 Redhat, pour créer des traitements à partir du référentiel d'actions disponibles dans le produit.
  - L'intégration d'une base de données depuis la version 8 permet d'augmenter la robustesse et la souplesse du produit par le stockage des traitements en base SQL. Cette base de données relationnelle stocke le paramétrage des workflows, les informations de traitement des flux métiers et l'historisation des tâches.
  - Un fichiers XML de configuration contient le paramétrage des différentes tâches implémentées pour chaque workflow configuré.
  - Une Interface Administrateur de gestion des flux permet le suivi et la relance des jobs.

- **Avantages**

- Optimisation de la gestion des processus d'échange en masse de documents ou de données
- Centralisation règles de gestion
- Gestion des erreurs et procédure de reprise
- Traçabilité des actions

# Schéma Général



GED IBM

## Transformation de format

- Enrichissement de données & extraction de métadonnées
- Signature Electronique cachet serveur
- Signature Transactionnelle
- Implémentation des règles métiers
- Connexion aux applications métiers
- Traitement PDF :
  - ✓ Extraction de données des PDF textes
  - ✓ Couper les PDF sur mots Clefs
  - ✓ Compter les pages des PDF
  - ✓ Ajout de Fond de page sur les PDF

## Contrôle de champs obligatoires

Détection des doublons

Mise à jour en masse de métadonnées

Envoi de Notifications

Interrogation d'IA pour extraction de données

N°Siret / N°TVA IntraCo

GED IBM

# Interface Administrateur

The screenshot displays the DSINFORMATIQUE administrator interface. The main section is titled "Tâches en cours" and shows a list of tasks. A modal window provides details for a specific task, and a statistics chart is also visible.

**Liste des tâches DSIPILOT: 14 opérations**  
Base de données : D:/server/DSIPILOT\_8\_3\_BEAUVALLET.db

Id	Type	Workflow Name	Workflow Task Name	Status	Actions
20220622-165932-736-BEAUVALLET-TYPE_SYSTEM-DSI1818BECF803_7704359	TYPE_SYSTEM	BEAUVALLET	PurgeWk	WAIT	[Search] [Edit]
20220622-165934-910-BEAUVALLET-TYPE_BATCH-DSI1818BED0381	TYPE_BATCH	BEAUVALLET	ParserPDF	READY	[Search] [Edit]
20220622-165935-113-BEAUVALLET-TYPE_BATCH-DSI1818BED044C	TYPE_BATCH	BEAUVALLET	END	WAIT	[Search] [Edit]
20220622-165935-190-BEAUVALLET-TYPE_BATCH-DSI1818BED0498	TYPE_BATCH	BEAUVALLET	ParserPDF	NO_CRITICAL	[Search] [Edit]
20220622-165935-190-BEAUVALLET-TYPE_BATCH-DSI1818BED0498	TYPE_BATCH	BEAUVALLET	END	END	[Search] [Edit]
20220622-165935-190-BEAUVALLET-TYPE_BATCH-DSI1818BED0498	TYPE_BATCH	BEAUVALLET	END	NO_CRITICAL	[Search] [Edit]
20220622-165935-190-BEAUVALLET-TYPE_BATCH-DSI1818BED0498	TYPE_BATCH	BEAUVALLET	END	READY	[Search] [Edit]
20220622-165935-190-BEAUVALLET-TYPE_BATCH-DSI1818BED0498	TYPE_BATCH	BEAUVALLET	END	NO_CRITICAL	[Search] [Edit]
20220622-165935-190-BEAUVALLET-TYPE_BATCH-DSI1818BED0498	TYPE_BATCH	BEAUVALLET	END	READY	[Search] [Edit]
20220622-165935-190-BEAUVALLET-TYPE_BATCH-DSI1818BED0498	TYPE_BATCH	BEAUVALLET	END	NO_CRITICAL	[Search] [Edit]
20220622-165935-190-BEAUVALLET-TYPE_BATCH-DSI1818BED0498	TYPE_BATCH	BEAUVALLET	END	READY	[Search] [Edit]
20220622-165935-190-BEAUVALLET-TYPE_BATCH-DSI1818BED0498	TYPE_BATCH	BEAUVALLET	END	NO_CRITICAL	[Search] [Edit]
20220622-165935-190-BEAUVALLET-TYPE_BATCH-DSI1818BED0498	TYPE_BATCH	BEAUVALLET	END	READY	[Search] [Edit]
20220622-165935-190-BEAUVALLET-TYPE_BATCH-DSI1818BED0498	TYPE_BATCH	BEAUVALLET	END	NO_CRITICAL	[Search] [Edit]
20220622-165935-190-BEAUVALLET-TYPE_BATCH-DSI1818BED0498	TYPE_BATCH	BEAUVALLET	END	READY	[Search] [Edit]

**Visualisation de la tâche: 20220622-165932-736-BEAUVALLET-TYPE\_SYSTEM-DSI1818BECF803\_7704359**

**Visualisation des statistiques**

**Visualisation**  
Analyse statistiques

Suivi dans le temps / lignes | Suivi dans le temps / bars | Affichages avec filtres

Semaine | Mois | Années

Graphique montrant l'évolution des statistiques (Entités success, Entités errors, Actions count, Actions errors, Actions success, Number of tasks) sur une période allant de 2022-05-24 à 2022-06-22.

- Description générale :
  - Nombre total d'opérations
  - Chemin d'accès à la base de données
- Dans le tableau les tâches sont affichées avec les informations suivantes :
  - Id de la tâche
  - Type de tâche
  - Nom du workflow
  - Non de la tâche dans ce workflow
  - Statut de la tâche
  - Actions disponibles
- Statistiques disponibles
  - Par critère d'entités
  - Par workflow
  - Par date
  - Il est possible de regrouper les données par période (mois ou année)

# Fonctionnalités

- Import des documents et métadonnées de différentes sources
  - Serveurs de Fichiers – Base de Données – CSV – XML – Json – EXCEL- Fichiers plats
  - Appels de WebServices métiers
  - Kofax – Docushare – Domino – IBM Content Manager – IBM CM Ondemand — SharePoint (Online / OnPremise)
- Modules Disponibles
  - Application de règles métiers
  - Extraction de données des PDF textes\*
  - Couper les PDF sur mots Clefs, Compter les pages des PDF\*
  - Ajout de Fond de page sur les PDF\*
  - Signature PDF (Cachet Serveur & Transactionnelle)\*
  - Création PDF Factur-X\*
  - Génération Flux XML UBL, XML CII\*
  - Contrôle des données et notamment N°Siret et TVA IntraCO via base INSEE\*
  - Appels aux IA IBM & Recital\*

\*en option

- Export vers des référentiels
  - IBM Content Manager – SharePoint (Online / OnPremise) – Messagerie POP / IMAP
  - Système de Fichier – Base de Données
  - Chorus-Pro avec gestion des statuts
  - Infor M3 – SAP Business One - Divalto – Sage - Etc... (*développement sur demande\**)
  - Génération de XML, CSV, JSON pour les Application Métiers
  - *Vers le PPF dès qu'il sera disponible (2024-2025)*
  
- Actions sur les référentiels
  - Import – Mise à Jour – Suppression – Export - Reclassement
  - Création de l'arborescence des dossiers
  - Récupération de l'ID des documents pour mise à disposition des ERP
  - Conversion AFP vers PDF pour Content Manager OnDemand\*
  - Détection des doublons

\*En option

- Lancement DSIPilot
  - Service Windows
  - Appel ligne de commande
  - Webservices
  
- Traçabilité
  - Chaque action est tracée dans la base de données



# Quelques références



Une équipe à votre écoute

Téléphone : +33 (0)4 42 53 01 81

Email : [contact@dsidiff.fr](mailto:contact@dsidiff.fr)

[www.dsidiff.fr](http://www.dsidiff.fr)