



Informatica®  
10.5.7

# Glossaire

© Copyright Informatica LLC 2016, 2025

Ce logiciel et la documentation associée sont fournis uniquement sous un accord de licence séparé contenant des restrictions d'utilisation et de divulgation. Il est interdit de reproduire ou de transmettre sous quelle que forme et par quel que moyen que ce soit (électronique, photocopie, enregistrement ou autre) tout ou partie de ce document sans le consentement préalable d'Informatica LLC.

Informatica et le logo Informatica sont des marques ou des marques déposées d'Informatica LLC aux États-Unis et dans de nombreux autres pays. La liste actuelle des marques de commerce de Informatica est disponible sur le Web à l'adresse <https://www.informatica.com/trademarks.html>. Les autres noms de société ou de produit peuvent être des marques de commerce ou des marques déposées de leurs détenteurs respectifs.

Soumis à vos droits de retrait, le logiciel transmettra automatiquement certaines informations à Informatica (aux États-Unis) concernant l'environnement informatique et réseau dans lequel le Logiciel est déployé et les statistiques du système et d'utilisation des données du déploiement. Cette transmission est considérée comme faisant partie des Services selon la politique de confidentialité d'Informatica et Informatica utilisera et traitera par ailleurs ces informations conformément à la politique de confidentialité d'Informatica disponible sur <https://www.informatica.com/in/privacy-policy.html>. Il est possible de désactiver la collecte d'utilisation dans l'outil Administrator tool.

**U.S. GOVERNMENT RIGHTS** Les programmes, les logiciels, les bases de données et les documents connexes et les données techniques fournis aux clients du gouvernement américain sont des « logiciels commerciaux » ou des « données techniques commerciales », conformément au règlement fédéral sur les acquisitions et aux règlements supplémentaires propres à l'Agence. En tant que tel, l'utilisation, la duplication, la divulgation, la modification et l'adaptation sont assujetties aux restrictions et aux conditions de licence énoncées dans le contrat gouvernemental applicable et, dans la mesure applicable par les termes du contrat gouvernemental, les droits additionnels énoncés dans la réglementation FAR 52.227-19, licence de logiciel d'ordinateur commercial.

Le produit inclut les logiciels ACE(TM) et TAO(TM), copyright Douglas C. Schmidt et son groupe de recherche à Washington University, University of California, Irvine, et Vanderbilt University, Copyright (©) 1993-2006, tous droits réservés.

Ce produit inclut le logiciel Curl, copyright 1996-2013, Daniel Stenberg, <daniel@haxx.se>. Tous droits réservés. Les autorisations et limitations concernant ce logiciel sont sujettes aux conditions publiées sur <http://curl.haxx.se/docs/copyright.html>. L'autorisation d'utiliser, copier, modifier et distribuer ce logiciel à toute fin, avec ou sans rémunération, est accordée par les présentes, à la condition que la notification de copyright ci-dessus et cette notification d'autorisation apparaissent dans toutes les copies. Le produit inclut des logiciels sous copyright 2001-2005 (©) MetaStuff, Ltd. Tous droits réservés. Les autorisations et limitations concernant ce logiciel sont sujettes aux conditions publiées sur <http://www.dom4j.org/license.html>.

Le produit inclut des logiciels sous copyright © 2004-2007, The Dojo Foundation. Tous droits réservés. Les autorisations et limitations concernant ce logiciel sont sujettes aux conditions publiées sur <http://dojotoolkit.org/license>.

Ce produit inclut le logiciel ICU sous copyright de International Business Machines Corporation et autres. Tous droits réservés. Les autorisations et limitations concernant ce logiciel sont sujettes aux conditions publiées sur <http://source.icu-project.org/repos/icu/icu/trunk/license.html>.

Ce produit inclut le logiciel OSSP UUID sous copyright © 2002 Ralf S. Engelschall, copyright © 2002 The OSSP Project Copyright © 2002 Cable & Wireless Deutschland. Les autorisations et limitations concernant ce logiciel sont sujettes aux conditions publiées sur <http://www.opensource.org/licenses/mit-license.php>.

Consultez les brevets applicables à l'adresse <https://www.informatica.com/legal/patents.html>.

Les renseignements contenus dans cette documentation sont sujets à modification sans préavis. Si vous constatez des problèmes liés à la documentation, merci de les signaler par courriel à l'adresse [infa\\_documentation@Informatica.com](mailto:infa_documentation@Informatica.com).

Les produits Informatica sont garantis conformément aux termes et conditions des accords en vertu desquels ils sont fournis. INFORMATICA FOURNIT LES INFORMATIONS DE CE DOCUMENT « EN L'ÉTAT » SANS GARANTIE D'AUCUNE SORTIE, EXPRESSE OU IMPLICITE, NOTAMMENT AUCUNE GARANTIE DE QUALITÉ MARCHANDE, D'ADAPTATION À UN USAGE PARTICULIER ET D'ABSENCE DE CONTREFAÇON

Certaines parties de ce logiciel et/ou de cette documentation sont soumises à des droits d'auteur détenus par des tiers. Les notifications de tiers requises sont incluses avec le produit.

Date de publication: 2025-02-26

# Sommaire

- Préface..... 4**
- Ressources Informatica. . . . . 4
  - Informatica Network. . . . . 4
  - Base de connaissances Informatica. . . . . 4
  - Documentation Informatica. . . . . 4
  - Matrices de disponibilité des produits Informatica. . . . . 5
  - Informatica Velocity. . . . . 5
  - Informatica Marketplace. . . . . 5
  - Support client international Informatica. . . . . 5
- Annexe A: Glossaire..... 6**

# Préface

Pour en savoir plus sur la terminologie des produits Informatica que vous pouvez utiliser avec Informatica Administrator, Informatica Analyst et Informatica Developer, consultez le *Glossaire Informatica®*.

## Ressources Informatica

Informatica vous fournit toute une gamme de ressources de produits via Informatica Network et autres portails en ligne. Utilisez ces ressources pour tirer le meilleur parti de vos produits et solutions Informatica, et pour apprendre d'autres utilisateurs et experts en la matière d'Informatica.

### Informatica Network

Informatica Network est la passerelle à de nombreuses ressources, y compris la base de connaissances Informatica et le support client international Informatica. Pour accéder à Informatica Network, visitez le site <https://network.informatica.com>.

En tant que membre d'Informatica Network, vous disposez des options suivantes :

- Rechercher les ressources de produits dans la base de connaissances.
- Afficher les informations de disponibilité des produits.
- Créer et vérifier vos dossiers de support.
- Rechercher votre réseau de groupe d'utilisateurs local Informatica et collaborer avec vos pairs.

### Base de connaissances Informatica

Utilisez la base de connaissances Informatica pour rechercher des ressources de produits telles que des articles pratiques, des meilleures pratiques, des didacticiels vidéo et des questions fréquemment posées.

Pour effectuer des recherches dans la base de connaissances, visitez le site <https://search.informatica.com>. N'hésitez pas à contacter l'équipe de la base de connaissances Informatica à l'adresse [KB\\_Feedback@informatica.com](mailto:KB_Feedback@informatica.com) pour lui faire part de vos questions, commentaires et suggestions concernant la base de connaissances.

### Documentation Informatica

Utilisez le portail de documentation Informatica pour explorer une vaste bibliothèque de documentation pour les versions de produits actuelles et récentes. Pour explorer le portail de documentation, visitez le site <https://docs.informatica.com>.

N'hésitez pas à contacter l'équipe Documentation Informatica à l'adresse [infa\\_documentation@informatica.com](mailto:infa_documentation@informatica.com) pour lui faire part de vos questions, commentaires ou suggestions concernant la documentation des produits.

## Matrices de disponibilité des produits Informatica

Les matrices de disponibilité des produits (PAM) indiquent les versions des systèmes d'exploitation, les bases de données et les types de source et cible de données pris en charge par une version d'un produit. Vous pouvez parcourir les PAM Informatica à l'adresse <https://network.informatica.com/community/informatica-network/product-availability-matrices>.

## Informatica Velocity

Informatica Velocity est un ensemble de conseils et de meilleures pratiques développés par les services professionnels d'Informatica et basés sur les expériences réelles de centaines de projets de gestion des données. Informatica Velocity représente le savoir collectif de consultants d'Informatica qui collaborent avec des organisations du monde entier pour planifier, développer, déployer et gérer des solutions performantes de gestion des données.

Vous trouverez les ressources d'Informatica Velocity à l'adresse <http://velocity.informatica.com>. Si vous avez des questions, des commentaires ou des suggestions sur Informatica Velocity, contactez les services professionnels d'Informatica à l'adresse [ips@informatica.com](mailto:ips@informatica.com).

## Informatica Marketplace

Informatica Marketplace est un forum dans lequel vous pouvez trouver des solutions qui permettent d'augmenter et d'améliorer vos implémentations Informatica. Exploitez les centaines de solutions de développeurs et de partenaires Informatica sur Marketplace pour améliorer votre productivité et accélérer le délai d'implémentation de vos projets. Vous trouverez Informatica Marketplace à l'adresse <https://marketplace.informatica.com>.

## Support client international Informatica

Vous pouvez contacter un centre de support international par téléphone ou via le réseau Informatica.

Pour rechercher le numéro de téléphone du support client international Informatica local, visitez le site Web Informatica à l'adresse <https://www.informatica.com/services-and-training/customer-success-services/contact-us.html>.

Pour trouver des ressources de support en ligne sur le réseau Informatica, visitez le site <https://network.informatica.com> et sélectionnez l'option eSupport.

# ANNEXE A

## Glossaire

### **adresse de courriel dynamique**

Adresse de courriel définie dans un paramètre ou une variable de flux de travail.

### **application**

Objet que vous pouvez déployer dans un service d'intégration de données. Il peut contenir des objets de données, des mappages, des services de données SQL, des services Web et des flux de travail.

### **application complète**

Une application que vous devez redéployer dans le service d'intégration de données après la modification des objets d'application.

### **application d'exécution**

Application déployée dans un service d'intégration de données et contenant des instances d'exécution d'objets d'application.

### **application en phase de conception**

Application que vous modifiez dans Developer tool. L'application contient des instances en phase de conception des objets d'application.

### **application incrémentielle**

Application que vous pouvez mettre à jour en déployant un correctif d'application dans le service d'intégration de données.

### **audit**

Ensemble de règles et de conditions qui valident les données à la source et à la cible d'un mappage. Les règles d'audit calculent une valeur agrégée pour une colonne de données dans la source ou la cible. Les conditions d'audit comparent plusieurs règles d'audit ou bien des règles d'audit et des valeurs constantes.

### **base de données virtuelle**

Un service de données SQL que vous pouvez interroger. Elle contient des objets virtuels et offre une vue uniforme des données provenant de sources de données distinctes et hétérogènes.

### **bibliothèque de définitions de types**

Objet du référentiel modèle utilisé pour le stockage des définitions de types de données complexes pour un mappage ou un mapplet.

**big data**

Un ensemble de données qui est si grand et si complexe qu'il ne peut pas être traité via des outils de gestion de base de données standard.

**Blaze Executor**

Composant du DTM qui peut simplifier et convertir un mappage en un plan d'exécution Blaze exécuté sur une grappe Hadoop.

**cible dynamique**

Fichier plat ou cible relationnelle d'un mappage qui peut changer lors de l'exécution. Les transformations Écriture peuvent définir les colonnes cibles lors de l'exécution en fonction du flux de mappage ou à partir d'une cible associée. Les transformations Écriture peuvent également abandonner et remplacer la table cible lors de l'exécution.

**clé candidate**

Colonne ou ensemble de colonnes qui identifie de façon unique chaque ligne source dans une table de base de données.

**clé documentée**

Clé primaire déclarée dans la base de données sources.

**clé inférée**

Colonne ou ensemble de colonnes que l'outil Analyst ou l'outil Developer considère comme une clé candidate en fonction des données de colonne.

**colonne dépendante**

Dans une dépendance fonctionnelle, la colonne contenant les valeurs qui sont déterminées par une colonne déterminante.

**colonne déterminante**

Dans une dépendance fonctionnelle, un ensemble de colonnes qui détermine la valeur de la colonne dépendante. Si la déterminante comporte zéro colonnes, la dépendante est une constante.

**composant de démarrage**

Composant exécutable démarré en premier par la Data Transformation lors de l'exécution d'une transformation Processeur de données.

**CompressionCodec**

Interface de compression Hadoop. Un codec est l'implémentation d'un algorithme de compression/décompression. Dans Hadoop, un codec est représenté par une implémentation de l'interface CompressionCodec.

**configuration de type**

Ensemble de propriétés de ports complexes qui spécifient le type de données des éléments de type de données complexes ou le schéma des données.

**conservation**

Processus de validation et de gestion des métadonnées d'une source de données découvertes afin qu'elles soient exploitables et puissent figurer dans des rapports.

**conteneur**

Allocation de mémoire et de ressources processeur sur un nœud doté du rôle de calcul. Un service d'application utilise le conteneur pour effectuer des calculs à distance sur le nœud. Par exemple, un service d'intégration de données qui s'exécute sur une grille peut exécuter un mappage à distance dans un conteneur sur un nœud doté du rôle de calcul.

**contenu de courriel dynamique**

Contenu d'un courriel défini dans un paramètre ou une variable de flux de travail.

**correctif d'application**

Objet que vous pouvez déployer dans un service d'intégration de données. Il contient les métadonnées que vous pouvez utiliser pour mettre à jour une application incrémentielle déployée.

**correspondance directe**

Dans une recherche globale, une correspondance directe est une ressource qui correspond à l'ensemble de la requête de recherche. Dans la recherche de découverte, une correspondance directe est une correspondance avec tout ou partie des métadonnées de la ressource qui correspond à la requête de recherche.

**correspondance indirecte**

Correspondance dans les résultats de recherche de découverte qui est liée à la ressource correspondant directement en tout ou en partie à la requête de recherche.

**DataNode**

Nœud HDFS qui stocke les données dans le système de fichiers Hadoop. Une grappe HDFS peut avoir plusieurs DataNode comprenant des données reproduites entre eux.

**découverte de clé étrangère**

Le processus qui recherche dans une source de données les colonnes correspondant aux colonnes de clé primaire dans la source de données parent.

**découverte de clé primaire**

Le processus d'identification d'une colonne ou d'une combinaison de colonnes qui identifie de façon unique une ligne dans la source de données.

**découverte de dépendance fonctionnelle**

Le processus qui recherche les relations de dépendance fonctionnelle entre les colonnes d'une source de données.

**découverte de domaines de données**

Processus qui identifie tous les domaines de données associés à une colonne en fonction des valeurs ou du nom de la colonne.

**découverte des données d'entreprise**

Processus qui recherche les statistiques de profil de colonne, les domaines de colonne, les clés principales et les clés étrangères dans un grand nombre de sources de données réparties sur plusieurs connexions ou schémas.

**définition de type de données complexes**

Représentation réutilisable du schéma des données que vous référencez dans un port de structure ou dans un port complexe qui contient des éléments de type structure. Un ou plusieurs ports complexes peuvent utiliser la définition de type de données complexes.

**définition de type de données imbriquées**

Définition de type de données complexes qui référence d'autres définitions de types de données complexes.

**définition de type de données récursives**

Définition de type de données imbriquées dans laquelle l'une des définitions de type de données complexes, quel que soit son niveau, est la même que celle d'un parent.

**dépendance fonctionnelle**

La relation existant entre les colonnes d'un ensemble de colonnes d'une table donnée, dans laquelle la colonne déterminante détermine la colonne dépendante de manière fonctionnelle.

**déployer**

Rendre les objets d'une application accessibles aux utilisateurs finaux. Selon les types d'objets dans l'application, les utilisateurs finaux peuvent ensuite exécuter des requêtes sur les objets, accéder aux services Web, exécuter des mappages ou des flux de travail.

**destinataire**

Utilisateur ou groupe dans le domaine Informatica qui reçoit une notification lors d'un flux de travail.

**destinataire dynamique**

Destinataire de notification défini dans un paramètre ou une variable de flux de travail.

**développement collaboratif**

Collaboration des membres de l'équipe sur un projet de développement. La collaboration inclut des fonctionnalités telles que le versionnage via l'extraction et l'archivage d'objets de référentiel.

**Distribution Cloudera, notamment Apache Hadoop (CDH)**

Version Cloudera de l'infrastructure du logiciel open-source Hadoop.

**document d'exemple de source**

Exemple de documents traité par une transformation Processeur de données.

**document de sortie**

Document résultant d'une transformation Processeur de données.

**document source**

Document d'entrée d'une transformation Processeur de données.

**domaine de données**

Objet de référentiel modèle prédéfini ou défini par l'utilisateur qui représente la signification fonctionnelle d'une colonne en fonction des données ou du nom de la colonne. Par exemple, le numéro de sécurité sociale, le numéro de carte de crédit ou l'identifiant de messagerie.

**données hiérarchiques**

Ensemble de données hiérarchiquement liées. La relation hiérarchique est représentée sous forme de schéma. Les transformations Informatica utilisent des types de données complexes pour représenter les données hiérarchiques.

**données virtuelles**

Les informations obtenues lorsque vous interrogez les tables virtuelles ou exécutez des procédures stockées dans un service de données SQL.

**Dossier**

Conteneur pour les objets du référentiel modèle. Utilisez les dossiers pour organiser les objets dans un projet et créer des dossiers pour grouper les objets en fonction des besoins de l'entreprise.

**entrée de tâche**

Données transmises dans une tâche par les variables et les paramètres de flux de travail. La tâche utilise les données d'entrée pour effectuer une unité de travail.

**entrepôt de profilage**

Base de données relationnelle qui stocke les informations de profilage, telles que les résultats de profils et de fiches d'évaluation.

**environnement d'exécution**

Environnement que vous configurez pour exécuter un mappage ou un profil. L'environnement d'exécution peut être Natif ou Hive.

**environnement de validation**

Environnement que vous configurez pour valider un mappage ou un profil. Vous validez un mappage ou un profil afin de vous assurer qu'il peut s'exécuter dans un environnement d'exécution. L'environnement de validation peut être Hive, natif ou les deux.

**environnement Hadoop**

Environnement que vous pouvez configurer pour exécuter un mappage ou un profil sur une grappe Hadoop. Vous devez configurer Hadoop en tant qu'environnement de validation et d'exécution.

**environnement natif**

Environnement par défaut dans le domaine Informatica qui exécute un mappage, un flux de travail ou un profil. Le service d'intégration de données effectue l'extraction, la transformation et le chargement de données. Ce service peut également transmettre le traitement à un environnement non natif, qui représente un cluster distribué en dehors du domaine Informatica, tel que Hadoop ou Databricks.

**environnement non natif**

Cluster distribué en dehors du domaine Informatica, tel que Hadoop ou Databricks, où le service d'intégration de données peut transmettre le traitement de l'exécution. Si le service d'intégration de données ne transmet pas le traitement à un environnement non natif, il traite la tâche dans un environnement natif du domaine Informatica.

**événement**

Objet de flux de travail qui démarre ou termine le flux de travail. Un événement représente quelque chose se produisant pendant l'exécution du flux de travail. L'éditeur affiche les événements sous forme de cercles.

**Événement du processeur de données**

Occurrence survenant pendant l'exécution d'une transformation Processeur de données.

**Exécuteur Hive**

Composant du DTM qui peut simplifier et convertir un mappage ou un profil en un plan d'exécution Hive exécuté sur une grappe Hadoop.

**exécution du gestionnaire de Data Transformation (DTM)**

Composant de calcul du service d'intégration de données qui extrait, transforme et charge les données de façon à terminer une tâche de Data Transformation.

**expression de prédicat**

Une expression qui filtre les données d'un mappage. Une expression de prédicat renvoie true ou false.

**fiche d'évaluation**

Représentation graphique de valeurs valides pour une colonne source ou sortie d'une règle dans les résultats de profil. Utilisez des fiches d'évaluation pour mesurer la progression de la qualité des données.

**Fichier schéma XSD**

Fichier\* .xsd contenant un schéma XML qui définit les éléments, les attributs et la structure de documents XML.

**flux de séquence**

Connecteur entre des objets de flux de travail spécifiant l'ordre d'exécution des objets par le service d'intégration de données. L'éditeur affiche les flux de séquence sous forme de flèches.

**flux de séquence conditionnel**

Flux de séquence qui comprend une expression à laquelle le service d'intégration de données donne la valeur True ou False. Si l'expression vaut True, le service d'intégration de données exécute l'objet suivant dans le flux de travail. Si l'expression vaut False, le service d'intégration de données n'exécute pas l'objet suivant dans le flux de travail.

**flux de séquence par défaut**

Flux de séquence sortant d'une passerelle exclusive qui donne toujours la valeur True. Lorsque tous les autres flux de séquence conditionnels donnent la valeur false, le service d'intégration de données exécute l'objet connecté au flux de séquence sortant par défaut.

**flux de travail**

Représentation graphique d'un ensemble d'événements, de tâches et de décisions qui définissent un processus métier. Utilisez l'outil Developer pour ajouter des objets à un flux de travail et pour connecter les objets aux flux de séquence. Le service d'intégration de données utilise les instructions configurées dans le flux de travail pour exécuter les objets.

**fonction complexe**

Type de fonction prédéfinie dans laquelle le type d'entrée ou de retour est un type de données complexes.

**gestionnaire de Data Transformation logique (LDTM)**

Composant du service d'intégration de données qui optimise et compile des tâches, puis les envoie au gestionnaire de Data Transformation d'exécution (DTM).

**glossaire de domaine de données**

Conteneur de tous les domaines de données et les groupes de domaine de données dans l'outil Analyst et l'outil Developer.

**Grappe Hadoop**

Grappe des machines configurée pour exécuter les applications et les services Hadoop. Une grappe Hadoop typique inclut un nœud principal et plusieurs nœuds de travail. Le nœud principal exécute les démons principaux JobTracker et NameNode. Un nœud de travail exécute les démons DataNode et TaskTracker. Dans les petits clusters, le nœud principal peut également exécuter les démons de travail.

**groupe de domaines de données**

Ensemble de domaines de données d'une catégorie de domaine de données spécifique.

**groupe de mesures**

Un groupe de mesures définies par l'utilisateur.

**Hadoop Distributed File System (HDFS)**

Système de stockage de fichiers distribué utilisé par des applications Hadoop.

**Hive**

Infrastructure d'entrepôt de données conçue pour d'Hadoop. Hive prend en charge un langage de type SQL appelé HiveQL pour le résumé de données, la requête et l'analyse.

**ID de l'instance de flux de travail**

Un nombre qui identifie de façon unique une instance de flux de travail exécutée.

**Informatica Administrator**

Informatica Administrator (l'outil Administrator) est une application qui consolide les tâches administratives pour les objets du domaine comme les services, les nœuds, les licences et les grilles. Vous gérez le domaine et la sécurité du domaine via l'outil Administrator.

**Informatica Developer**

Informatica Developer (l'outil Developer) est une application qui vous permet de concevoir des solutions d'intégration de données. Le référentiel modèle stocke les objets créés dans l'outil Developer.

## **Informatica Mass Ingestion**

Informatica Mass Ingestion (outil Mass Ingestion) est une application que vous pouvez utiliser pour configurer, déployer, exécuter et surveiller les spécifications d'ingestion.

## **ingestion de masse**

Déplacement de grandes quantités de données entre une base de données relationnelle et un cluster Data Lake ou Hadoop.

## **instance de flux de travail**

Représentation d'exécution d'un flux de travail. Quand vous exécutez un flux de travail depuis une application déployée, vous exécutez une instance du flux de travail. Vous pouvez exécuter plusieurs instances du même flux de travail simultanément.

## **instance DTM**

Représentation logique spécifique de l'exécution du gestionnaire de Data Transformation (DTM) que le service d'intégration de données crée pour exécuter une tâche. Selon la configuration du service d'intégration de données, les instances DTM peuvent s'exécuter dans le processus de service d'intégration de données, dans un processus DTM distinct sur le nœud local ou dans un processus DTM distinct sur un nœud distant.

## **Intelligent Structure Discovery**

Service d'Informatica Intelligent Cloud Services qui utilise les algorithmes d'apprentissage automatique pour identifier automatiquement les types d'information dans un fichier d'entrée. Intelligent Structure Discovery crée une structure intelligente, un modèle de structure et de types de champ de données découvert dans le fichier. Le modèle inclut des éléments simples et complexes, selon l'organisation des données dans le fichier.

## **JobTracker**

Service Hadoop qui coordonne les tâches de mappage et de réduction et les planifie pour qu'elles s'exécutent sur TaskTrackers.

## **lien d'exécution**

Lien de groupe à groupe qui utilise une stratégie et/ou un paramètre afin de déterminer les ports à lier entre les groupes lors de l'exécution.

## **lignage des fiches d'évaluation**

Diagramme qui affiche l'origine des données, décrit le chemin et indique comment les données circulent pour une métrologie ou un groupe de métrologies dans une fiche d'évaluation. Dans l'analyse du lignage des fiches d'évaluation, les zones ou nœuds représentent des objets. Les flèches représentent les relations des flux de données.

## **magasin de métadonnées**

Base de données utilisée par Hive pour stocker les métadonnées des tables Hive stockées dans HDFS. Les magasins de métadonnées peuvent être locaux, intégrés ou distants.

## **mappage**

Un ensemble d'entrées et de sorties liées par des objets de transformation qui définissent les règles de transformation de données.

**mappage d'écriture d'objet de données logique**

Un mappage qui enregistre des données dans des cibles en utilisant un objet de données logiques comme entrée. Il contient un ou plusieurs objets de données logiques comme entrée et un objet de données physiques comme cible.

**mappage d'objet de données logique**

Un mappage qui lie un objet de données logiques à un ou plusieurs objets de données physiques. Il peut comprendre une logique de transformation.

**mappage d'opération**

Un mappage qui effectue l'opération de service Web pour le client du service Web. Un mappage d'opération peut contenir une transformation Entrée, une transformation Sortie et plusieurs transformations Erreur.

**mappage de grille**

Mappage Informatica que le moteur Blaze compile et distribue sur une grappe de nœuds.

**mappage de lecture d'objet de données logique**

Un mappage qui fournit une vue des données via un objet de données logiques. Il contient un ou plusieurs objets de données physiques comme sources et un objet de données logiques comme sortie de mappage.

**mappage de table virtuelle**

Un mappage qui contient une table virtuelle comme cible.

**mappage dynamique**

Mappage dans lequel vous pouvez modifier les sources, les cibles et la logique de transformation lors de l'exécution en fonction des paramètres et des règles que vous définissez. Vous pouvez configurer des mappages dynamiques afin d'autoriser les modifications de métadonnées sur les sources et les cibles. Vous pouvez déterminer les ports qu'une transformation reçoit, ceux à utiliser dans la logique de transformation ainsi que les liens à établir entre les groupes de transformation.

**mapplet**

Un objet réutilisable qui contient un ensemble de transformations que vous pouvez utiliser pour plusieurs mappages ou valider comme règle.

**MapReduce**

Modèle de programmation pour le traitement de grands volumes de données en parallèle.

**mesure**

Colonne d'une source de données ou de sortie d'une règle faisant partie d'une fiche d'évaluation.

**mise en cache de l'ensemble des résultats**

Un cache qui contient les résultats de chaque demande de service de données SQL ou de service Web. Avec la mise en cache des ensembles de résultats, le service d'intégration de données renvoie les résultats en cache lorsque les utilisateurs exécutent des requêtes identiques. La mise en cache des ensembles de résultats diminue le temps d'exécution des requêtes similaires.

**modèle d'objet de données logique**

Un modèle de données qui décrit les données dans une entreprise et la relation entre ces données. Il contient des objets de données logiques et définit les relations entre eux.

**module de service SQL**

Service de composant du service d'intégration de données qui gère les requêtes SQL envoyées à un service de données SQL depuis un outil client tiers.

**module de service Web**

Un composant dans le service d'intégration de données qui gère les requêtes des opérations de services Web envoyées à un service Web à partir d'un client de service Web.

**Module Service de flux de travail**

Composant du service d'intégration de données gérant les demandes d'exécution de flux de travail.

**NameNode**

Nœud de la grappe Hadoop qui gère l'espace de nom du système de fichiers et entretient l'arborescence du système de fichiers ainsi que les métadonnées de tous les fichiers et les répertoires dans l'arborescence.

**nœud**

Une représentation d'un niveau dans la hiérarchie d'un message de service Web.

**objet de données logique**

Un objet qui décrit une entité logique dans une entreprise. Il possède des attributs, des clés et il décrit les relations qui existent entre les attributs.

**objet de données personnalisé**

Un objet de données physiques utilisant une ou plusieurs ressources relationnelles ou objets de données relationnels liés comme sources. Un objet de données personnalisé permet de réaliser des tâches comme une jointure de données à partir de ressources liées ou de lignes de filtre. L'objet de données personnalisé utilise une connexion unique et une instruction SQL pour les tables sources.

**objet de données physiques**

Une représentation physique de données utilisée pour lire, rechercher ou écrire sur des ressources.

**opérateur complexe**

Type d'opérateur permettant de référencer des noms d'éléments ou d'accéder à des éléments dans un type de données complexes.

**opération d'écriture d'objet de données**

Objet de référentiel contenant les propriétés requises pour effectuer certaines opérations d'exécution sur des cibles. Une opération d'écriture d'objet de données est associée à un objet de données cible.

**opération de lecture d'objet de données**

Objet de référentiel contenant les propriétés requises pour effectuer certaines opérations d'exécution sur des sources. Une opération de lecture d'objet de données est associée à un objet de données source.

### **optimisation basée sur le coût**

Une méthode d'optimisation qui réduit le temps d'exécution des mappages qui réalisent des opérations de jointure. Avec l'optimisation basée sur le coût, le service d'intégration de données crée différents plans d'exécution d'un mappage et calcule un coût pour chacun des plans. Le service d'intégration de données exécute le plan présentant le coût le plus faible. Le service d'intégration de données calcule le coût selon des statistiques de la base de données, les E/S, le processeur, le réseau et la mémoire.

### **optimisation de projection précoce**

Une méthode d'optimisation qui réduit le volume de données transféré entre les transformations dans le mappage. Avec l'optimisation de projection précoce, le service d'intégration de données identifie les ports inutilisés et supprime les liens entre les ports d'un mappage.

### **optimisation de sélection précoce**

Une méthode d'optimisation qui réduit le nombre de lignes qui passent par le mappage. Avec l'optimisation de sélection précoce, le service d'intégration de données rapproche les filtres de la source du mappage dans le pipeline.

### **optimisation de semi-jointure**

Une méthode d'optimisation qui réduit le nombre de lignes extraites à partir de la source. Avec l'optimisation de semi-jointure, le service d'intégration de données modifie les opérations de jointure d'un mappage. Le service d'intégration de données applique la méthode d'optimisation de semi-jointure à une transformation Jointure lorsqu'un groupe d'entrée plus grand comporte des lignes qui ne correspondent pas à un petit groupe d'entrée de la condition de jointure. Le service d'intégration de données lit les lignes à partir du plus petit groupe, trouve les lignes correspondantes dans le groupe plus grand puis effectue l'opération de jointure.

### **optimisation du prédicat**

Une méthode d'optimisation qui simplifie ou réécrit les expressions de prédicat dans un mappage. Avec l'optimisation de prédicat, le service d'intégration de données tente d'appliquer des expressions de prédicat le plus tôt possible pour améliorer la performance de mappage.

### **optimisation du refoulement**

Une méthode d'optimisation qui pousse la logique de transformation vers une base de données sources ou cibles. Avec l'optimisation du refoulement, le service d'intégration de données traduit la logique de transformation en requêtes SQL et les envoie à la base de données. La base de données exécute les requêtes SQL pour traiter les données.

### **Outil Informatica Monitoring**

L'outil Informatica Monitoring tool (l'outil Monitoring tool) est une application qui fournit un lien direct vers l'onglet Surveiller de l'outil Administrator tool. L'onglet Surveiller affiche les propriétés, les statistiques d'exécution et les rapports d'exécution relatifs aux objets d'intégration exécutés dans un service d'intégration de données.

### **paramètre de flux de travail**

Valeur constante que vous définissez avant l'exécution du flux de travail. Les paramètres conservent la même valeur pendant toute la durée d'exécution du flux de travail. Vous définissez la valeur du paramètre dans un fichier de paramètres. Tous les paramètres de flux de travail sont définis par l'utilisateur.

## **partitionnement**

Processus de division des données sous-jacentes en sous-ensembles pouvant être exécutés dans plusieurs threads de traitement. Lorsque les administrateurs activent le service d'intégration de données pour optimiser le parallélisme, le service augmente le nombre de threads de traitement, ce qui peut optimiser les performances de mappage et de profilage.

## **passerelle**

Objet de flux de travail qui fractionne et fusionne les chemins dans le flux de travail en fonction de la manière dont le service d'intégration de données évalue les expressions dans les flux de séquence conditionnels. L'éditeur affiche les passerelles sous forme de losanges.

## **Passerelle exclusive**

Passerelle qui représente une décision effectuée dans un flux de travail. Lorsqu'une passerelle exclusive fractionne le flux de travail, le service d'intégration de données effectue une décision afin de prendre l'une des branches sortantes. Lorsqu'une passerelle exclusive fusionne le flux de travail, le service d'intégration de données attend l'issue d'une branche entrante avant de déclencher la branche sortante.

## **pipeline**

Source et toutes les transformations et les cibles qui reçoivent les données de cette source. Chaque mappage contient un ou plusieurs pipelines.

## **Plan d'exécution Hive**

Série de tâches Hive qui l'exécuteur Hive génère une fois qu'il a traité un mappage ou un profil. Un plan d'exécution Hive peut également être appelé flux de travail Hive.

## **Poids métrique**

Nombre entier supérieur ou égal à 0 affecté à une mesure. Un poids métrique définit la contribution de la mesure au score du groupe de mesures.

## **point de partition**

Limite entre les étapes d'un pipeline de mappage. Lorsque le partitionnement est activé, le service d'intégration de données peut redistribuer les lignes de données aux points de partition.

## **port complexe**

Type de port auquel est attribué un type de données complexes, tel qu'un tableau, une structure ou une carte, pour transmettre des données hiérarchiques.

## **port complexe imbriqué**

Port complexe contenant une définition de type de données complexes imbriquées.

## **port dynamique**

Port qui peut recevoir une ou plusieurs colonnes d'une transformation en amont et créer un port généré pour chaque colonne.

## **port généré**

Port dans un port dynamique qui représente une seule colonne. L'outil Developer tool génère des ports en fonction d'une ou de plusieurs règles d'entrée.

**Port variable d'état**

Port variable qui fait référence aux valeurs des lignes précédentes.

**préprocesseur**

Processeur de document utilisé pour effectuer une modification globale sur un document source préalablement à la transformation principale.

**procédure stockée virtuelle**

Un ensemble d'instructions procédurales ou de flux de données dans un service de données SQL.

**processeur de document**

Composant agissant sur un document dans son entièreté et effectuant le plus souvent des conversions préliminaires à l'analyse.

**processus DTM**

Processus de système d'exploitation que le service d'intégration de données démarre pour exécuter les instances DTM. En fonction de la manière dont vous configurez le service d'intégration de données, il peut exécuter chaque instance DTM dans un processus DTM distinct sur le nœud local ou sur un nœud distant.

**profil**

Un objet contenant des règles régissant la découverte de formes dans les données sources. Exécutez un profil pour évaluer la structure des données et vérifier que les colonnes de données contiennent les types d'informations que vous prévoyez.

**profil d'objet de données**

Un objet du référentiel qui définit le type d'analyse que vous effectuez sur la source de données.

**profil de colonne**

Type de profil qui détermine les caractéristiques des colonnes dans une source de données, telles que la fréquence de la valeur, les pourcentages, les modèles et les types de données.

**profil de découverte des données d'entreprise**

Type de profil que vous utilisez pour effectuer la découverte des données d'entreprise.

**profil de jointure**

Un type de profil qui détermine le degré de chevauchement entre un ensemble composé d'une ou plusieurs colonnes dans une source de données et un ensemble similaire dans la même source de données ou une source de données différente.

**profil de système d'exploitation**

Type de sécurité que les services d'intégration de données sous UNIX ou Linux utilisent pour isoler les environnements d'exécution. Le profil de système d'exploitation contient le nom d'utilisateur du système d'exploitation, les variables du processus de service, les variables d'environnement et les autorisations. Le service d'intégration de données exécute les mappages, les flux de travail, les tâches de profilage et les fiches d'évaluation à l'aide des autorisations de l'utilisateur du système d'exploitation et des propriétés définies dans le profil de système d'exploitation.

## **Projet**

Conteneur de niveau supérieur qui stocke les objets dans Informatica Analyst et Informatica Developer. Créez des projets en fonction d'objectifs ou de besoins professionnels. Les projets s'affichent dans Informatica Analyst et Informatica Developer.

## **recherche de découverte**

Type de recherche dans l'outil Analyst qui identifie les ressources en fonction de correspondances directes avec la requête de recherche et de relations avec d'autres objets qui correspondent à la requête de recherche.

## **règle d'entrée**

Règle qui détermine les ports générés à créer dans un port dynamique.

## **règle de données**

Logique métier réutilisable qui identifie une colonne par ses valeurs comme appartenant à un domaine de données particulier.

## **règle de nom de colonne**

Logique métier réutilisable qui identifie une colonne par son nom comme appartenant à un domaine de données particulier.

## **règle de profilage**

Logique métier réutilisable qui définit les conditions appliquées aux données lorsque vous exécutez un profil. Utilisez les règles pour valider les données dans un profil et mesurer la progression de la qualité des données. Vous pouvez créer une règle dans Informatica Analyst ou Informatica Developer.

## **rôle d'utilisateur**

Ensemble de privilèges que vous attribuez à un utilisateur ou à un groupe. Vous attribuez des rôles à des utilisateurs et des groupes pour le domaine et pour certains services d'application du domaine.

## **rôle du nœud**

Objectif d'un nœud. Un nœud doté du rôle de service peut exécuter des services d'application. Un nœud doté du rôle de calcul peut effectuer les calculs requis par des services d'application distante. Un nœud doté des deux rôles peut exécuter des services d'application et effectuer localement des calculs pour ces services.

## **schéma**

Définition de la structure de données. Les ports complexes de type de données de structure utilisent des définitions de type de données complexes pour représenter le schéma.

## **schéma virtuel**

Un schéma dans une base de données virtuelle qui définit la structure de la base de données.

## **Schéma XML**

Définition des éléments, des attributs et de la structure utilisés dans des documents XML. Le schéma est conforme à la norme de schéma XML World Wide Web Consortium et est stocké en tant que fichier \*.xsd.

**score de groupe de mesures**

Moyenne pondérée calculée de tous les scores de mesures du groupe de mesures.

**score de mesure**

Pourcentage de valeurs valides dans une mesure.

**Scripts Hive**

Script en langage de requête Hive qui contient des requêtes et des commandes Hive permettant d'exécuter le mappage.

**segment de grille**

Partie d'un mappage de grille qui est contenue dans une tâche de grille.

**service d'application**

Service qui s'exécute sur un ou plusieurs nœuds dans le domaine Informatica. Vous pouvez créer et gérer les services d'application dans Informatica Administrator ou via le programme de commande infacmd. Les services d'application incluent des services qui peuvent avoir plusieurs instances dans le domaine et des services système qui peuvent n'en avoir qu'une seule. Configurez chaque service d'application en fonction de vos exigences d'environnement.

**Service d'ingestion de masse**

Service d'application qui gère les spécifications d'ingestion de masse que vous créez dans l'outil Ingestion de masse. Le service d'ingestion de masse valide les spécifications, les spécifications de planifications à exécuter sur un service d'intégration de données et surveille les statistiques d'ingestion.

**Service d'intégration de données**

Service d'application qui effectue des tâches d'intégration de données pour Informatica Analyst, Informatica Developer et les clients externes. Les tâches d'intégration de données incluent l'aperçu des données et l'exécution des mappages, des profils, des services de données SQL, des services Web et des flux de travail.

**service de données**

Un ensemble d'opérations réutilisables que vous pouvez exécuter pour accéder aux données et les transformer. Un service de données fournit un modèle unifié de données auquel vous pouvez accéder par le biais d'un service Web ou par rapport auquel vous pouvez exécuter une requête SQL.

**service de données SQL**

Une base de données virtuelle que vous pouvez interroger. Elle contient des objets virtuels et offre une vue uniforme des données provenant de sources de données distinctes et hétérogènes.

**Service de gestionnaire de ressource**

Service qui gère des ressources de calcul dans le domaine et répartit les tâches afin d'obtenir des performances et une évolutivité optimales. Le service de gestionnaire de ressource collecte des informations sur les nœuds dotés du rôle de calcul. Le service fait correspondre les spécifications de la tâche avec la disponibilité des ressources afin d'identifier le nœud de calcul le plus approprié pour exécuter la tâche. Le service de gestionnaire de ressource communique avec les nœuds de calcul dans une grille du service d'intégration de données. Activez le service de gestionnaire de ressource lorsque vous configurez une grille du service d'intégration de données pour exécuter les tâches dans des processus distants distincts.

### **Service de référentiel modèle**

Un service d'application dans le domaine Informatica qui exécute et gère le référentiel modèle. Le référentiel modèle stocke des métadonnées créées par des produits Informatica dans une base de données relationnelle pour permettre la collaboration entre les produits.

### **service système**

Service d'application qui peut présenter une seule instance dans le domaine. Les services système sont créés lorsque vous créez le domaine. Vous pouvez activer, désactiver et configurer les services système.

### **sortie de tâche**

Données transmises d'une tâche vers des variables de flux de travail. Lorsque vous configurez une tâche, spécifiez les valeurs de sortie de la tâche que vous voulez affecter aux variables de flux de travail. Le service d'intégration de données copie les valeurs de sortie de la tâche dans les variables de flux de travail à l'issue de la tâche. Le service d'intégration de données peut accéder à ces valeurs depuis les variables du flux de travail quand il évalue des expressions dans des flux de séquence conditionnels et lorsqu'il exécute d'autres objets dans le flux de travail.

### **source dynamique**

Fichier plat ou source relationnelle d'un mappage qui peut changer lors de l'exécution. Les transformations Lecture et Recherche peuvent obtenir la définition ou les modifications de métadonnées directement depuis la source. Si vous utilisez un paramètre pour la source, vous pouvez modifier la source lors de l'exécution.

### **Sparkline**

Graphique en courbes qui affiche la variation d'une valeur Null, d'une valeur unique ou d'une valeur non-unique au cours des cinq dernières exécutions consécutives de profil.

### **spécification d'ingestion de masse**

Configuration qui détermine le mode d'ingestion des données source d'une base de données relationnelle à un emplacement spécifique d'un cluster Data Lake ou Hadoop.

### **structure intelligente**

Modèle de structure et de types de champ de données qu'Intelligent Structure Discovery identifie dans un fichier d'entrée. Le modèle inclut des éléments simples et complexes, selon l'organisation des données dans le fichier.

### **table virtuelle**

Une table d'une base de données virtuelle.

### **tâche**

Objet de flux de travail qui exécute une seule unité de travail dans le flux de travail, comme l'exécution d'un mappage, l'envoi d'un courriel ou l'exécution d'une commande Shell. Une tâche représente quelque chose qui est exécuté pendant le flux de travail. L'éditeur affiche les tâches sous forme de carrés.

### **tâche de commande**

Tâche de pré ou post-traitement de données locales d'un flux de travail de moteur Blaze.

### **tâche de grille**

Demande de tâche de traitement parallèle. Lorsque le mappage s'exécute dans l'environnement Hadoop, l'exécuteur du moteur Blaze envoie la demande au gestionnaire de grille. Lorsque le mappage s'exécute dans l'environnement natif et que le service d'intégration de données s'exécute en mode distant, le processus du service d'intégration de données envoie la demande au gestionnaire de service sur le nœud de calcul principal.

### **Tâche Hive**

Tâche du plan d'exécution Hive. Un plan d'exécution Hive contient plusieurs tâches Hive. Une tâche Hive contient un script Hive.

### **Tâche MapReduce**

Unité de travail qui comprend des données d'entrée, le programme MapReduce et les informations de configuration. Hadoop exécute la tâche MapReduce en la divisant en tâches de mappage et de réduction.

### **tasklet**

Partition d'un segment de grille qui s'exécute sur un autre DTM.

### **TaskTracker**

Nœud dans la grappe Hadoop qui exécute les tâches telles que celles de mappage ou de réduction. Les TaskTrackers envoient les rapports de progression au JobTracker.

### **transformation**

Un objet de référentiel dans un mappage qui génère, modifie ou transmet des données. Chaque transformation a une fonction différente.

### **transformation du service Web**

Une transformation qui traite les requêtes ou les réponses de service Web. Parmi les transformations de service Web, on trouve la transformation Entrée, la transformation Sortie, la transformation Erreur et la transformation Consommateur de service Web.

### **type de données complexes**

Un type de données qui permet de représenter plusieurs valeurs de données dans une seule position de colonne. Les valeurs de données sont appelées éléments.

### **type de données imbriquées**

Type de données complexes contenant au moins un élément d'un type de données complexes. Par exemple, un type de données de structure qui contient un élément de type tableau.

### **type de données primitives**

Type de données qui permet de représenter une valeur de données unique dans une seule position de colonne.

### **Valeur hors norme**

Une valeur hors norme est un modèle, une valeur ou une fréquence dans une colonne des résultats de profil qui ne se situe pas dans la plage de valeurs attendue.

**variable de flux de travail**

Valeur qui peut être modifiée lors de l'exécution d'un flux de travail. Utilisez des variables de flux de travail pour référencer des valeurs et enregistrer des informations d'exécution. Vous pouvez utiliser des variables de flux de travail système ou définies par l'utilisateur.

**variable de flux de travail/définie par l'utilisateur**

Variable de flux de travail qui capture la sortie de tâche ou qui capture les critères que vous indiquez. Après avoir créé une variable de flux de travail définie par l'utilisateur, vous configurez le flux de travail pour affecter une valeur d'exécution à la variable.

**variable système de flux de travail**

Variable de flux de travail qui renvoie des informations d'exécution système, comme l'identifiant d'instance de flux de travail, l'utilisateur ayant démarré le flux de travail ou la date de début du flux de travail.

**vue virtuelle des données**

Une base de données virtuelle définie par un service de données SQL que vous pouvez interroger comme s'il s'agissait d'une base de données physique.

**XMap**

Objet de transformation Processeur de données qui mappe un document d'entrée XML sur un autre document XML.

**XPath**

Langage de requête utilisé pour sélectionner les nœuds dans un document XML et effectuer des calculs.