

Spectrum™ Technology Platform

Version 12.0 SP1

Guide Metadata Insights



Table des matières

1 - Démarrage

Compréhension des ressources de données de votre entreprise	4
Découverte de Metadata Insights	5
Utilisations communes	8
Connexion aux données	9
Exemple de modèle	64

2 - Modeling

Logical Model	71
Modèle physique	74
Mapping	82
Model Store	98

3 - Profilage

Création d'un profil	112
Spécification des paramètres de profilage par défaut	123
Analyse d'un profil	124
Affichage de rapports d'analyse du profil	129

4 - Généalogie et analyse de l'impact

Affichage de la généalogie et de l'analyse de l'impact	133
Icônes dans Généalogie et Analyse de l'impact	135
Utilise	140

1 - Démarrage

In this section

Compréhension des ressources de données de votre entreprise	4
Découverte de Metadata Insights	5
Utilisations communes	8
Connexion aux données	9
Exemple de modèle	64

Compréhension des ressources de données de votre entreprise

Si votre entreprise est comme la plupart des entreprises, vous disposez d'un grand nombre de ressources de données contenant tout, des coordonnées des clients et de leur historique d'achats aux données financières, en passant par les enregistrements des transactions, etc. Ces systèmes peuvent être exécutés sur plateformes différentes, parfois gérées par des services différents avec des contrôles de sécurité différents. Il existe potentiellement une multitude de données à votre disposition pour répondre aux questions commerciales. Cependant, il est difficile de savoir quels systèmes contiennent les données dont vous avez besoin, quels systèmes sont fiables et comment ils sont tous connectés.

Metadata Insights offre la visibilité qu'il vous faut pour identifier les sources de données les plus fiables à utiliser pour répondre à une demande commerciale.

1. Commencez par connecter les actifs de données physiques de votre entreprise à Spectrum™ Technology Platform. Reportez-vous à la section [Data Source Connections](#) à la page 9.
2. Ensuite, définissez un modèle de données physique pour représenter vos ressources de données dans Metadata Insights. Ce processus vous permet de comprendre comment vos ressources de données sont structurées, par exemple, les tables et les colonnes de chaque base de données et les relations entre les tables. Reportez-vous à la section [Ajout d'un modèle de données physique](#) à la page 74.
3. Une fois que vous connaissez les ressources de données physiques qui sont à votre disposition, vous voudrez vous assurer que les données sous-jacentes sont de bonne qualité. Utilisez le profilage pour analyser vos ressources de données, déterminer les types de données qu'elles contiennent (tels que les noms, les adresses électroniques et les devises) et identifier les données incomplètes et non conformes. Reportez-vous à la section [Création d'un profil](#).

Conseil : En utilisant les rapports du profilage, vous pouvez créer des flux Spectrum™ Technology Platform afin d'améliorer la qualité des données. Si vous ne disposez pas de licence pour l'un des modules de qualité des données Spectrum™ Technology Platform, contactez votre responsable de compte Pitney Bowes.

4. Grâce à la création d'un modèle de données physique et à une bonne compréhension de l'état de vos données via le profilage, vous pouvez créer des Logical Models pour représenter les entités commerciales que votre entreprise souhaite comprendre, telles que des clients, des fournisseurs ou des produits. Dans cette procédure, vous sélectionnez les sources des données que vous souhaitez utiliser pour renseigner chaque entité, telles que les adresses des clients et l'historique des achats. Reportez-vous à la section [Création d'un modèle logique](#) à la page 71.
5. Pour mettre à jour vos ressources de données, vous devez comprendre la manière dont elles sont toutes connectées et comment les données fluent d'une source vers une destination. Utilisez la fonction Analyse de généalogie et d'impact de Metadata Insights pour afficher les

dépendances entre les sources de données, les destinations et les processus qui utilisent les données. Grâce à ces informations, vous pouvez prendre des décisions éclairées sur l'impact d'une modification apportée aux sources de données, résoudre des résultats inattendus et comprendre comment des entités Spectrum™ Technology Platform telles que des flux, des sous-flux et des bases de données Spectrum s'affectent mutuellement. Pour plus d'informations, reportez-vous à la section [Affichage de la généalogie et de l'analyse de l'impact](#) à la page 133.

Découverte de Metadata Insights

Metadata Insights vous donne le contrôle dont vous avez besoin pour fournir à votre entreprise des informations tirées de données précises et dans les temps impartis. Utilisez Metadata Insights pour développer des modèles de données, afficher le flux de données de la source vers l'application commerciale et évaluer la qualité de vos données grâce au profilage. Grâce à ces informations, vous pouvez identifier les ressources de données à utiliser pour répondre à des questions professionnelles spécifiques, adapter et optimiser les processus afin d'améliorer l'utilité et la cohérence des données au sein de l'entreprise toute entière et résoudre les problèmes de données.

Pour accéder à Metadata Insights, ouvrez un navigateur Web et accédez à :

```
http://server:port/metadata-insights
```

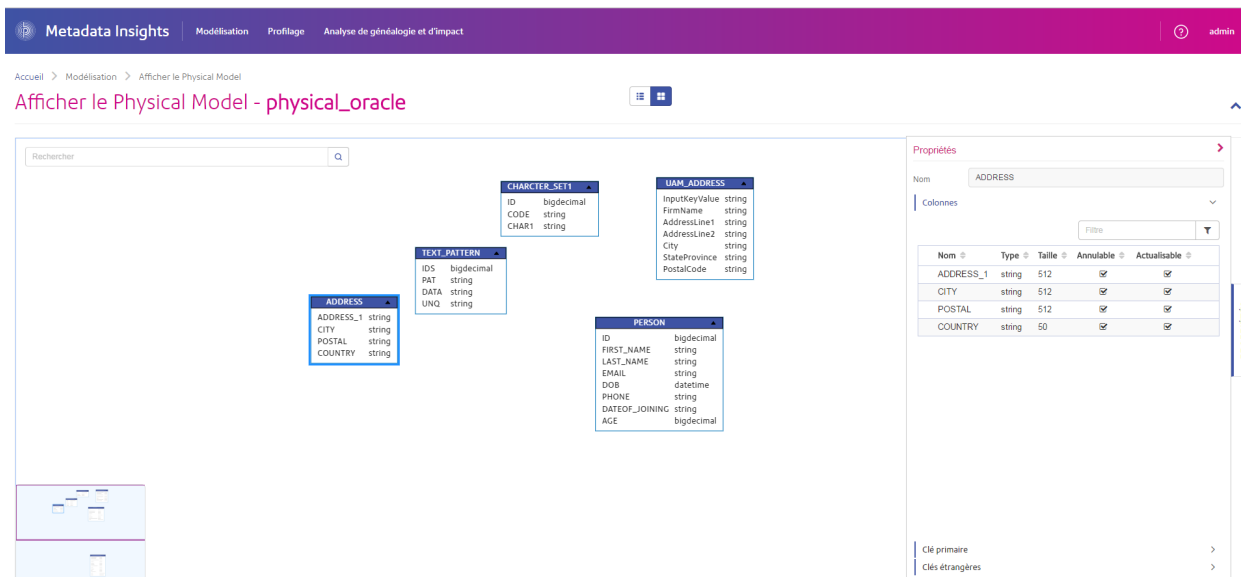
Où : *server* est l'adresse IP ou le nom de serveur de votre serveur Spectrum™ Technology Platform et *port* le port HTTP. Par défaut, le port HTTP est 8080.

Les fonctions de Metadata Insights sont divisées dans les domaines suivants : modeling, mappage, profilage et généalogie et analyse de l'impact.

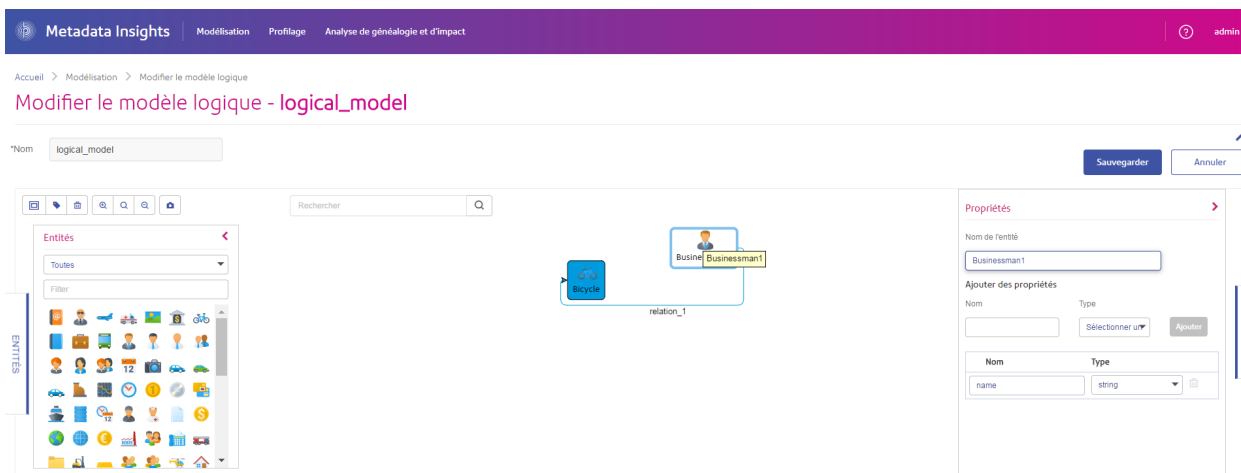
Modeling

La vue Modeling est celle où vous créez des modèles de données physiques et logiques et où vous les déployez dans un Model Store, créant ainsi une couche d'abstraction sur les sources de données sous-jacentes du serveur Spectrum™ Technology Platform.

Un *Physical Model* organise les actifs de données de votre entreprise de manière significative. Un Physical Model permet d'extraire les données à partir de tables, de colonnes et de vues individuelles pour créer une ressource unique que vous pouvez ensuite utiliser pour fournir des données à des Logical Models ou pour effectuer un profilage.



Un *Logical Model* définit les objets qui intéressent votre entreprise et les attributs de ces objets, ainsi que la manière dont ces objets sont reliés entre eux. Par exemple, un Logical Model pour un client peut contenir des attributs de nom et de date de naissance. Il peut également avoir une relation avec un objet d'adresse de domicile, qui contient des attributs de lignes d'adresse, de ville et de code postal. Une fois que vous avez défini les attributs des objets qui intéressent votre entreprise, vous pouvez mapper des sources de données physiques aux attributs du Logical Model, améliorant ainsi l'identification des actifs de données spécifiques qui seront utilisés pour renseigner cet attribut.



Profilage

La prise de décisions commerciales éclairées nécessite des données de qualité. C'est pourquoi il est important pour vous d'avoir confiance dans l'exhaustivité, l'exactitude et la validité de vos données. Les enregistrements incomplets, les champs non conformes et le manque de contexte peuvent entraîner la fourniture de données trompeuses ou inappropriées à vos utilisateurs commerciaux, ce qui peut aboutir à des décisions faussées.

Le profilage des données peut vous donner confiance dans vos données. Le profilage analyse vos données et génère des rapports qui identifient les problèmes en matière d'exactitude, d'exhaustivité et de validité. Grâce à ces rapports, vous pouvez entreprendre des actions pour corriger les données incorrectes ou non conformes.

Metadata Insights fournit des outils de profilage qui permettent d'exécuter un profilage sur vos actifs de données, ainsi que sur les données qui alimentent les Logical Models et les Physical Models définis dans Metadata Insights. Grâce à ces informations, vous pouvez déterminer la fiabilité de vos données, créer des règles de qualité des données et appliquer des routines de standardisation et de normalisation pour résoudre les problèmes de qualité des données.

Metadata Insights | Modélisation | Profilage | Analyse de généalogie et d'impact | admin

Accueil > Profilage

Résultats de profilage des données

Nom de profil: logical_profile | Description: mssql_flatfile_oracle
 Démarré: 2017/03/30 8:43:10 PM | Source des données: logical_model
 Utilisateur: admin | Tables: 3
 Type de source de données: Logical | Enregistrements: 1672

Source de données: logical_model

Résumé

Complétude

98.50%

Incomplete Rows
Complete Rows

Résumé des tables

Nom de la table	Nombre de colonnes	Nombre d'enregistrements	Complétude(%)
Businessman1	1	1000	97.60
Bicycle	1	663	99.85
relation_1	1	9	100.00

Affichage de 3 enregistrement(s) sur 3 | Lignes par page: 10

Généalogie et analyse de l'impact

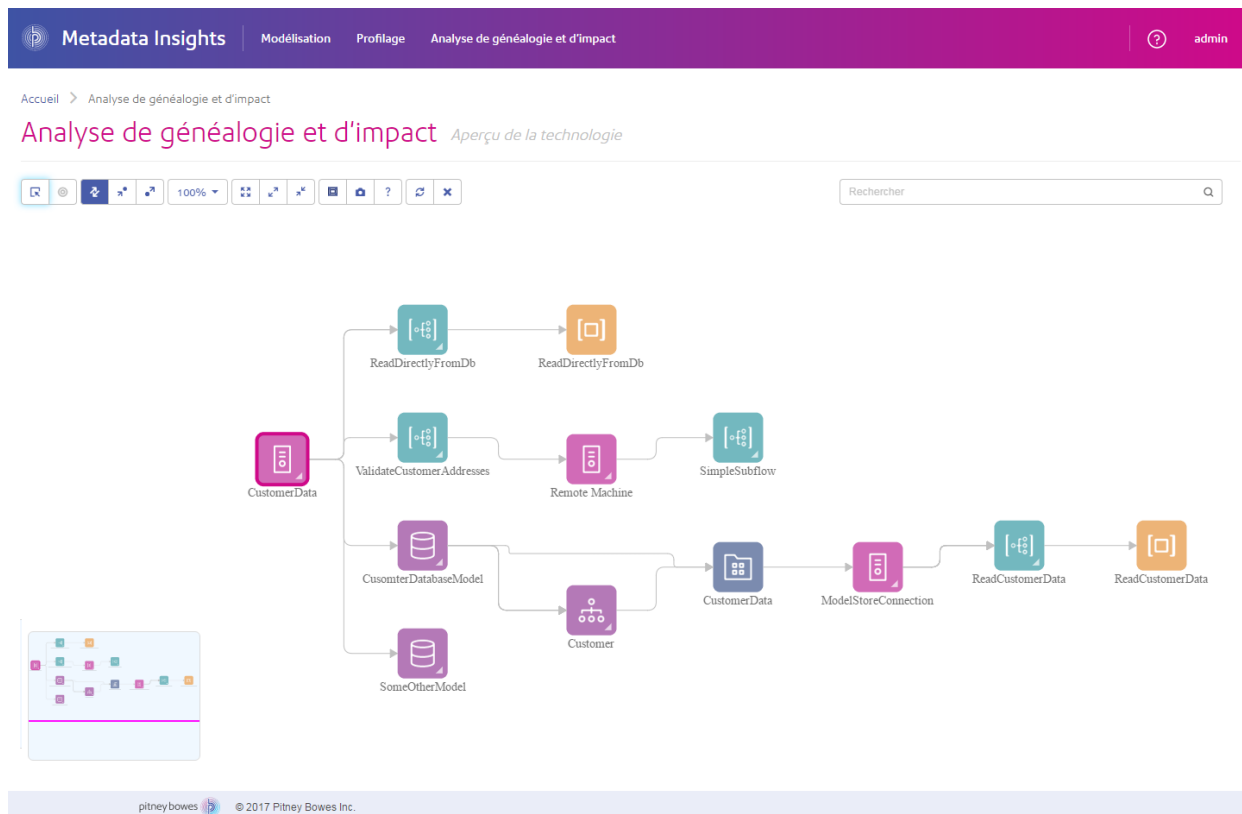
La vue Analyse de généalogie et d'impact affiche la manière dont les données fluent des sources de données vers les destinations de données et à travers les flux Spectrum™ Technology Platform. La généalogie et l'analyse de l'impact sont des concepts similaires qui décrivent différentes manières de tracer le flux de données.

Généalogie indique la provenance des données. Vous pouvez l'utiliser pour tracer le chemin d'accès aux données depuis leur source, en affichant tous les systèmes qui traitent et stockent les données en cours de route, tels que les flux Spectrum™ Technology Platform, les bases de données et les fichiers.

Analyse d'impact indique la destination des données et les systèmes qui dépendent des données d'une ressource de données sélectionnée. Vous pouvez l'utiliser pour afficher les flux, les bases de données et les fichiers qui utilisent une ressource de données directement ou indirectement.

L'analyse de l'impact s'avère utile si vous souhaitez comprendre comment une modification d'une base de données, d'un fichier ou d'un flux affecte les processus et les systèmes qui utilisent les données.

Metadata Insights peut afficher la généalogie et l'analyse de l'impact sur un seul diagramme qui montre le flux de données complet, de la source à la destination. Vous pouvez également choisir d'afficher uniquement la généalogie ou l'impact. En affichant ensemble la généalogie et l'analyse de l'impact des données, vous pouvez repérer les problèmes de vos processus de données et planifier des mises à niveau et des modifications à apporter à ces processus.



Utilisations communes

Évaluation de la qualité de vos données

Afin de fournir à votre entreprise des données fiables sur lesquelles baser ses décisions commerciales, vous devez être certain que vos données sont complètes et exemptes de toute erreur. Les enregistrements incomplets, les champs non conformes et le manque de contexte

peuvent entraîner la fourniture de données trompeuses ou inappropriées à vos utilisateurs commerciaux, ce qui peut aboutir à des décisions faussées.

Metadata Insights met à votre disposition des outils de profilage qui vous permettent d'évaluer la qualité des données contenues dans les ressources de données de votre entreprise. Vous pouvez utiliser le profilage pour déterminer la qualité générale des données d'une ou de plusieurs ressources de données et la qualité des données qui alimentent un modèle de données logique. Grâce à ces informations, vous pouvez déterminer la fiabilité de vos données, créer des règles de qualité des données et appliquer des routines de standardisation et de normalisation pour résoudre les problèmes de qualité des données. Ainsi, vous avez l'assurance que les données que vous fournissez à votre entreprise sont complètes et exemptes de toute erreur.

Pour plus d'informations, reportez-vous à la section [#unique_9](#).

Création de modèles de données logiques

Planification d'une mise à niveau de base de données

Connexion aux données

Data Source Connections

Une source de données est une base de données, un serveur de fichiers, un service Cloud ou une autre source de données que vous souhaitez traiter via Spectrum™ Technology Platform. Spectrum™ Technology Platform peut se connecter à plus de 20 types de sources de données.

Pour connecter Spectrum™ Technology Platform à une source de données, utilisez Management Console pour définir la connexion. Par exemple, si vous souhaitez lire les données d'un fichier XML dans un flux de données, et que le fichier XML est situé sur un serveur de fichiers distant, vous devez définir une connexion au serveur de fichiers avant de pouvoir définir le fichier XML d'entrée dans un flux de données. De même, pour écrire une sortie de flux de données dans une base de données, vous devez définir au préalable la base de données en tant que ressource externe.

Remarque : Pour lire ou écrire des données situées dans un fichier sur le serveur Spectrum™ Technology Platform lui-même, il n'est pas nécessaire de définir une connexion.

Prise en charge de la compression pour les serveurs de fichiers Cloud

Les serveurs de fichiers Amazon S3, Google cloud storage et MS Azure Blobstore prennent en charge les formats compressés `gzip (.gz)` et `zip (.zip)`.

Spectrum™ Technology Platform gère la compression et la décompression des fichiers écrits et lus vers/depuis les serveurs de fichiers.

Remarque : Vous pouvez utiliser le même serveur de fichiers pour gérer la lecture et l'écriture de fichiers normales et la compression et la décompression de fichiers.

Lecture d'un fichier au format compressé

Lors de la lecture d'un fichier à partir du serveur, son format de compression est calculé à partir de la propriété de clé de métadonnées `Content-Encoding` provenant du serveur.

Écriture d'un fichier au format compressé

Lors de l'écriture d'un fichier dans un serveur, indiquez le format de compression requis : `.gz` ou `.zip`. Le fichier est compressé en fonction de l'extension de compression spécifiée.

La propriété de clé de métadonnées `Content-Encoding` est également définie en fonction du format de compression sélectionné. Cette valeur de propriété est transmise au serveur de fichiers Cloud lors de l'écriture du fichier.

Entités et opérations prises en charge

Spectrum™ Technology Platform prend en charge ces entités et ces opérations pour chaque type de connexion :

Type de connexion	Propriétés de la table
-------------------	------------------------

Marketo	Les entités sont des types suivants :
---------	---------------------------------------

1. Entité
2. Entity Update

Remarque : Entity Update est une table virtuelle utilisée pour la mise à jour de l'entité prospect. Par exemple, **Merge_Leads** doit être utilisé pour la fusion des différents prospects Marketo.

MS Dynamics CRM	Les entités sont des types suivants :
-----------------	---------------------------------------

1. Appartenant à l'utilisateur
2. Appartenant à l'organisation
3. Appartenant à l'entreprise
4. Aucun

Type de connexion	Propriétés de la table
Netsuite	<p>Les entités sont des types suivants :</p> <ul style="list-style-type: none">• Enregistrements standard• Enregistrements personnalisés• Jointures• Recherches enregistrées <p>Lors de l'affichage du schéma Baseview créé, pour chaque type d'entité, le schéma résultant s'affiche.</p> <p>Par exemple, pour un enregistrement Recherche enregistrée, le schéma des résultats de recherche est affiché. Pour un enregistrement Jointure, le schéma des résultats de la jointure est affiché.</p> <p>Remarque : Dans une table de connexion NetSuite, la colonne de clé primaire est <code>internalId</code>.</p>
SAP	<p>1. Les colonnes d'entités sont de deux types :</p> <ul style="list-style-type: none">• Natif : les colonnes avec des types de données natifs sont affichées avec leurs types de données respectifs.• Personnalisé : les colonnes avec les types de données personnalisés sont affichées avec un type de données vide. <p>Pour déployer une Virtual Data Source calculé à partir d'une connexion SAP, assurez-vous que ses Metaviews et ses Baseviews incluent uniquement les entités dont les colonnes sont de types de données natifs. Si les Baseviews et les Metaviews disposent d'entités de types de données personnalisés, la Virtual Data Source ne peut pas être déployée.</p>
Siebel	<p>Les composants commerciaux sont affichés au format <code>Business Object.Business Component</code>.</p>

Type de connexion Propriétés de la table

Splunk

Opérations prises en charge	LIKE, ORDER BY, LIMIT, IN, BETWEEN, !=, <=, >=, <, >, différents opérateurs AND/OR	
Fonctions prises en charge	String Functions	upper, lower, length, len, ltrim, rtrim, substring, max, min
	Fonctions mathématiques	abs, ceil, exp, floor, sqrt, round

Remarque : Pour toutes les autres opérations de requête, utilisez la colonne `Splunk search` comme indiqué ci-dessous.

Spectrum™ Technology Platform fournit une colonne `search` dans la table `Splunk`, à l'aide de laquelle vous pouvez rechercher les données requises dans la connexion `Splunk`.

Lors de l'exécution d'une requête `select` sur `SplunkTable`, utilisez la colonne `search` dans la clause `where` dans l'un des scénarios ci-dessous :

1. Pour inclure les critères de recherche qui ne peuvent pas être spécifiés à l'aide de la syntaxe SQL ANSI.
2. Pour inclure les critères de recherche propres à Splunk qui ne peuvent pas être inclus dans le cadre de la requête SQL principale.

Par exemple, la requête ci-dessous recherche une valeur `_raw` qui contient la clé `opp` avec la valeur `ACC`.

```
select "_raw" from SplunkTable where "search"='search opp=ACC'
```

Type de connexion	Propriétés de la table
SuccessFactors	<p>Les entités sont de deux types :</p> <ol style="list-style-type: none"> 1. Entité : représente une table représentant une entité métier. 2. Jointure : Représente une correspondance entre deux tables de type <i>Entité</i> : une table parent et n'importe laquelle de ses tables enfant. <p>Remarque : Les liens ne sont pas présents entre les tables dans le schéma Baseview calculé à partir d'une connexion SuccessFactors. C'est parce que les clés étrangères ne figurent pas dans les tables SuccessFactors et que les jointures entre les tables sont indiquées par des tables de type <i>Jointure</i> dans la Spectrum™ Technology Platform.</p> <p>Les fonctions de tables de jointure sont les suivantes :</p> <ol style="list-style-type: none"> 1. Le nom d'une table de <i>jointure</i> indique que les deux tables <i>Entité</i> ont été mappées entre elles. 2. Chaque enregistrement d'une table de <i>jointure</i> contient la clé principale de l'entité parente et les colonnes de l'entité enfant respective. Par conséquent, la clé principale de l'entité parente particulière est mappée sur les détails de l'entité enfant. <p>Par exemple, <code>User#HR</code> est une table de <i>jointure</i> dans laquelle <code>User</code> est l'entité parente et <code>Hr</code> est l'entité enfant. Cette jointure représente tous les utilisateurs et leurs représentants HR respectifs. La table de jointure <code>User#HR</code> utilise par conséquent la table parent <code>User</code> dont la clé principale <code>UserId</code> est mappée vers les colonnes de la table enfant HR, comme <code>hr_userId</code>, <code>hr_username</code>, <code>hr_email</code>, etc.</p> <ol style="list-style-type: none"> 3. En cas de tables de <i>Jointure</i>, les fonctions <code>insert</code> et <code>update</code> s'utilisent de la même façon que la fonction <code>upsert</code>. C'est parce que les tables de <i>jointure</i> ne sont pas des entités réelles dans SuccessFactors, mais des mappages entre les entités et leurs propriétés de navigation ou leurs tables enfant. <p>Pour <code>insert/update</code> une table de <i>jointure</i> table, l'entité parente est mise à jour, tandis qu'un nouvel enregistrement est inséré dans la table enfant ou l'enregistrement existant est mis à jour d'après l'enregistrement parent.</p> <p>Remarque : Lors de la mise à jour, le mappage entre un parent et un enfant est modifié. Il est également possible de modifier les attributs individuels de l'enfant selon ce qui est requis.</p>
SugarCRM	<p>Opérations prises en charge :</p> <p><code>LIKE</code> (son opération est limitée à la sélection des options commençant par la valeur spécifiée, comme dans l'instruction <code>WHERE name LIKE 's%'</code>, qui sélectionne tous les noms commençant par la lettre S de l'alphabet), <code>ISNULL</code>, <code>IS NOT NULL</code>, <code>IN,NOT IN,>,>=,<,<=,=<></code>, <code>AND,OR</code></p>

Type de connexion	Propriétés de la table
Oracle Eloqua	<p>Opérations prises en charge</p> <p>Trois types d'entité sont pris en charge :</p> <ol style="list-style-type: none"> 1. Entité : indique une table représentant une entité métier. 2. Activité : indique une table représentant une entité métier dont les données sont générées en fonction d'une activité. 3. Entité personnalisée : indique les entités qui sont utilisées dans le cadre d'opérations spéciales fournies avec le Connecteur.


Types de connexion

Connexion à Amazon

Connexion à Amazon DynamoDB

Afin que Spectrum™ Technology Platform puisse accéder aux données d'Amazon DynamoDB, vous devez définir une connexion à Amazon DynamoDB à l'aide de Management Console. Une fois que cela est fait, vous pouvez créer des flux dans Enterprise Designer qui peuvent lire les données d'Amazon DynamoDB et y écrire des données.

Remarque : Cette connexion est à utiliser dans le module Metadata Insights.

1. Ouvrez Management Console.
2. Accédez à **Ressources > Sources de données**.
3. Cliquez sur le bouton Ajouter .
4. Dans le champ **Nom**, entrez un nom pour la nouvelle connexion. Ce nom est entièrement de votre choix.

Remarque : Une fois que vous enregistrez une connexion, vous ne pouvez plus modifier le nom.


5. Dans le champ **Type**, sélectionnez **Amazon DynamoDB**.
6. Dans le champ **ID de clé d'accès**, saisissez la séquence alphanumérique de 20 caractères qui vous a été fournie pour accéder à votre compte Amazon AWS.
7. Dans le champ **Clé d'accès secrète**, saisissez la clé de 40 caractères nécessaire pour authentifier la connexion.
8. Dans le champ **Région** sélectionnez la région du compte Amazon AWS.
9. Pour tester la connexion, cliquez sur **Tester**.
10. Cliquez sur **Enregistrer**.

Limitations Amazon DynamoDB

1. Les types de données hiérarchiques comme les listes, les cartes et les ensembles sont interprétés comme des types de données Chaîne. Cela est dû au fait que ces types de données ne sont pas pris en charge.
2. Les valeurs Null d'une source de données DynamoDB sont interprétées comme des valeurs de colonne vides.
3. La fonction d'agrégation `count` n'est pas prise en charge dans une requête sur un Model Store.

Connexion à Amazon S3

Afin que Spectrum™ Technology Platform puisse accéder aux données d'Amazon S3, vous devez définir une connexion à Amazon S3 à l'aide de Management Console. Une fois que cela est fait, vous pouvez créer des flux dans Enterprise Designer qui peuvent lire les données d'Amazon S3 et y écrire des données.

1. Ouvrez Management Console.
2. Accédez à **Ressources > Sources de données**.
3. Cliquez sur le bouton Ajouter .
4. Dans le champ **Nom**, entrez un nom pour la nouvelle connexion. Ce nom est entièrement de votre choix.

Remarque : Une fois que vous enregistrez une connexion, vous ne pouvez plus modifier le nom.

5. Dans le champ **Type**, sélectionnez **Cloud**.
6. Dans le champ **Service Cloud**, sélectionnez **AmazonS3**.
7. Dans le champ **Nom de compartiment**, saisissez le nom de compartiment tel que défini dans votre service Cloud Amazon S3. Il s'agit du compartiment où Spectrum™ Technology Platform lira et écrira des fichiers.
8. Saisissez votre clé d'accès et la clé secrète qui vous sont affectées par Amazon.
9. Dans le champ **Type de stockage**, sélectionnez le niveau de redondance que vous souhaitez autoriser pour le stockage des données.

Standard Niveau de redondance par défaut fourni par Amazon S3.

Redondance réduite Stocke les données non critiques et facilement reproduisibles à des niveaux de redondance inférieurs. Cela fournit un stockage assez fiable à moindre coût.

10. Dans la section **Chiffrement**, sélectionnez la méthode de chiffrement des données. Vous pouvez sélectionner le chiffrement côté serveur et/ou le chiffrement côté client.

Clé côté serveur Les données sont chiffrées et déchiffrées côté serveur. Vos données sont transmises sous forme de texte en clair au service Cloud Amazon, où elles sont chiffrées et stockées. Lors de leur récupération, les données sont déchiffrées

par le service Cloud Amazon, puis transmises sous forme de texte en clair à votre système.

Vous disposez de deux options permettant de spécifier la clé :

- **Géré par AWS** : la clé est automatiquement générée par le service Cloud Amazon S3.
- **Fourni par le client** : saisissez la clé à utiliser par le service Cloud Amazon S3 pour chiffrer et déchiffrer les données côté serveur.

Clé côté client Les données sont chiffrées et déchiffrées côté client. Les données sont chiffrées localement sur votre système client, puis transmises au stockage Cloud Amazon S3. Lors de leur récupération, les données sont retransmises à votre système au format chiffré et déchiffrées sur le système client.


Clé côté client : saisissez la clé à utiliser par votre système client pour chiffrer et déchiffrer les données.

Si vous sélectionnez à la fois **Clé côté serveur** et **Clé côté client**, le chiffrement et le déchiffrement sont effectués côté serveur et côté client. Les données sont tout d'abord chiffrées à l'aide de votre clé côté client, puis transmises au format chiffré à Amazon, où elles sont de nouveau chiffrées à l'aide de la clé côté serveur et stockées. Lors de leur récupération, Amazon commence par déchiffrer les données à l'aide de la clé côté serveur, puis transmet les données au format chiffré à votre système, où elles sont enfin déchiffrées à l'aide de la clé côté client.

Remarque : Pour utiliser la fonctionnalité de chiffrement du Cloud Amazon S3, vous devez installer les fichiers JAR Amazon S3 Security. Pour plus d'informations, reportez-vous à la section [Utilisation du chiffrement Cloud Amazon S3](#) à la page 17.

Pour plus d'informations sur les fonctionnalités de chiffrement d'Amazon S3, reportez-vous à l'adresse suivante :

docs.aws.amazon.com/AmazonS3/latest/dev/UsingEncryption.html

11. Si vous souhaitez définir des droits d'accès; dans la section **Droits d'accès**, cliquez sur .

Les trois types de bénéficiaires sont les suivants :

Everyone	Tous les utilisateurs, à l'exception des AuthenticatedUsers et du groupe LogDelivery.
AuthenticatedUsers	Pour les utilisateurs qui sont connectés à Amazon.
LogDelivery	Pour les utilisateurs qui écrivent des journaux d'activités dans un compartiment spécifié par l'utilisateur, si la journalisation de compartiment est activée.

Pour chaque bénéficiaire, sélectionnez les autorisations souhaitées :

Open/Download	L'utilisateur peut télécharger le fichier.
View	L'utilisateur peut afficher les autorisations existantes sur le fichier.

Edit

L'utilisateur peut modifier et définir les autorisations sur le fichier.

12. Pour tester la connexion, cliquez sur **Tester**.
13. Cliquez sur **Enregistrer**.

Utilisation du chiffrement Cloud Amazon S3

Pour utiliser la fonctionnalité de sécurité par chiffrement du service Cloud Amazon S3, vous devez télécharger les fichiers JAR de sécurité et les placer sur le serveur Spectrum™ Technology Platform. L'utilisation du chiffrement est facultatif.

1. Accédez au site de téléchargement.

Pour les plates-formes Windows et Linux utilisant Java7, les fichiers JAR peuvent être téléchargés depuis :

<http://www.oracle.com/technetwork/java/javase/downloads/jce-7-download-432124.html>

Pour les plates-formes AIX utilisant Java7, les fichiers JAR peuvent être téléchargés depuis :

<https://www14.software.ibm.com/webapp/iwm/web/preLogin.do?source=jcesdk>

2. Téléchargez ces deux fichiers JAR :

- local_policy.jar
- US_export_policy.jar

3. Placez ces fichiers JAR à l'emplacement suivant :


```
SpectrumFolder\Pitney Bowes\Spectrum\java64\jre\lib\security
```

4. Redémarrez le serveur.

Connexion à Amazon SimpleDB

Afin que Spectrum™ Technology Platform puisse accéder aux données d'Amazon SimpleDB, vous devez définir une connexion à Amazon SimpleDB à l'aide de Management Console. Une fois que cela est fait, vous pouvez créer des flux dans Enterprise Designer qui peuvent lire les données d'Amazon SimpleDB et y écrire des données.

Remarque : Cette connexion est à utiliser dans le module Metadata Insights.

1. Ouvrez Management Console.
2. Accédez à **Ressources > Sources de données**.
3. Cliquez sur le bouton Ajouter .
4. Dans le champ **Nom**, entrez un nom pour la nouvelle connexion. Ce nom est entièrement de votre choix.

Remarque : Une fois que vous enregistrez une connexion, vous ne pouvez plus modifier le nom.

5. Dans le champ **Type**, sélectionnez **Amazon SimpleDB**.
6. Dans le champ **ID de clé d'accès**, saisissez la séquence alphanumérique de 20 caractères qui vous a été fournie pour accéder à votre compte Amazon AWS.
7. Dans le champ **Clé d'accès secrète**, saisissez la clé de 40 caractères nécessaire pour authentifier la connexion.
8. Pour tester la connexion, cliquez sur **Tester**.
9. Cliquez sur **Enregistrer**.

Restrictions d'Amazon SimpleDB

Restriction d'écriture

Dans le stage Write to DB, le mode d'écriture **Mettre à jour** n'est pas disponible lors de l'écriture dans une table Amazon SimpleDB. L'option **Insérer** gère à la fois les opérations d'insertion et de mise à jour. Elle fait la distinction entre une insertion et une mise à jour à l'aide de la valeur unique de la colonne `ItemName` qui est présente dans toutes les tables Amazon SimpleDB.

Motif : une requête de mise à jour nécessite une clé principale pour chaque enregistrement de la table à mettre à jour, ce qui n'est pas pris en charge par les bases de données Amazon SimpleDB.


Restriction de lecture

Les fonctions d'agrégation `SUM` et `AVG` ne sont pas prises en charge lors de l'exécution des requêtes sur un Model Store.

Connexion à Apache Cassandra

Afin que Spectrum™ Technology Platform puisse accéder aux données d'une base de données Cassandra, vous devez définir une connexion à la base de données Cassandra à l'aide de Management Console. Une fois que cela est fait, vous pouvez créer des flux dans Enterprise Designer qui peuvent lire les données d'une base de données Cassandra et y écrire des données.

Remarque : Cette connexion est à utiliser dans le module Metadata Insights.

1. Ouvrez Management Console.
2. Accédez à **Ressources > Sources de données**.
3. Cliquez sur le bouton Ajouter .
4. Dans le champ **Nom**, entrez un nom pour la nouvelle connexion. Ce nom est entièrement de votre choix.

Remarque : Une fois que vous enregistrez une connexion, vous ne pouvez plus modifier le nom.

5. Dans le champ **Type**, sélectionnez **Apache Cassandra**.
6. Dans le champ **Hôte**, saisissez le nom de l'ordinateur ou l'adresse IP d'installation de la base de données Apache Cassandra.


7. Dans le champ **Espace de clé**, saisissez le nom de l'espace de clé du centre de données auquel vous souhaitez accéder.
8. Dans le champ **Port**, saisissez le port de configuration de la base de données Apache Cassandra.
9. Saisissez le nom d'utilisateur et le mot de passe à utiliser pour l'authentification auprès de la base de données Cassandra.
10. Dans le champ **Niveau de cohérence**, sélectionnez le niveau de cohérence des lignes de données sur les nœuds répliqués pour permettre une transaction de données réussie. Il peut s'agir d'un nœud, de tous les nœuds ou d'une combinaison de nœuds disponibles.
11. Dans **Taille de l'extraction**, saisissez le nombre de lignes de jeux de résultats que vous souhaitez extraire à chaque opération de lecture.
12. Pour tester la connexion, cliquez sur **Tester**.
13. Cliquez sur **Enregistrer**.

Restriction Apache Cassandra

La fonction d'agrégation `count` n'est pas prise en charge dans une requête sur un Model Store.

Connexion au Cloud Azure

Afin que Spectrum™ Technology Platform puisse accéder aux données de Microsoft Azure, vous devez définir une connexion à Microsoft Azure à l'aide de Management Console. Une fois que cela est fait, vous pouvez créer des flux dans Enterprise Designer qui peuvent lire les données de Microsoft Azure et y écrire des données.

1. Ouvrez Management Console.
2. Accédez à **Ressources > Sources de données**.
3. Cliquez sur le bouton Ajouter .
4. Dans le champ **Nom**, entrez un nom pour la nouvelle connexion. Ce nom est entièrement de votre choix.

Remarque : Une fois que vous enregistrez une connexion, vous ne pouvez plus modifier le nom.


5. Dans le champ **Type**, sélectionnez **Cloud**.
6. Dans le champ **Service Cloud**, sélectionnez **AzureBlobStorage**.
7. Dans le champ **Protocole**, sélectionnez si vous souhaitez que la connexion entre Azure et Spectrum™ Technology Platform utilise HTTP ou HTTPS.
8. Dans le champ **Nom de compte**, saisissez le nom de votre compte de stockage Azure.
9. Dans le champ **Clé d'accès** saisissez la clé d'accès à votre compte Azure.
10. Pour tester la connexion Cloud, cliquez sur **Tester**.
11. Cliquez sur **Enregistrer**.

Connexion à Flat File

Connexion à Delimited Flat File

Pour ajouter une nouvelle connexion Delimited Flat File, accédez à **Connexions > Flat File** et définissez **Type d'enregistrement** sur **Delimited**. Saisissez les détails d'accès du fichier et les détails du type de contenu pour permettre au module Data Federation de le lire correctement.

Remarque : Cette connexion est à utiliser dans le module Metadata Insights.


1. Accédez à **Connexions > Flat File**.
2. Par défaut, l'écran s'ouvre en mode Création. Sinon, cliquez sur  pour ajouter une nouvelle connexion Flat File.
3. Saisissez un **Nom de connexion** pour la connexion aux données Flat File.
4. Saisissez le **Chemin d'accès** en cliquant sur **Parcourir**, puis en sélectionnant le répertoire du fichier.
5. Sélectionnez le **Chiffrement de caractères** du fichier plat dans la liste déroulante.
6. Définissez le **Type d'enregistrement** sur **Delimited**.
7. Dans **Délimiteur de champ**, sélectionnez le séparateur prévu entre deux champs d'un enregistrement de fichier.
8. Sélectionnez le **Qualificatif de texte (facultatif)**, le cas échéant, qui englobe les valeurs de champ d'un enregistrement de fichier.
9. Dans **Séparateur de ligne**, la valeur **Par défaut** est sélectionnée, indiquant que le séparateur de ligne prévu dépend de l'exécution de Spectrum™ Technology Platform sous un système Unix ou Windows.
10. Pour indiquer si la première ligne du fichier est ou non une ligne d'en-tête, déplacez le curseur **La première ligne est l'enregistrement d'en-tête** sur **Oui** ou sur **Non**.
11. Pour spécifier si le type de données des différents champs d'un enregistrement du fichier doit être automatiquement détecté ou non, déplacez le curseur **Détecter le type de données du fichier** sur **Oui** ou sur **Non**.
12. Pour ignorer les enregistrements non conformes lors de l'analyse de fichier, déplacez le curseur **Ignorer les enregistrements incorrects** sur **Activé**.
13. Cliquez sur **Test**.
Un message confirme que le test de la connexion a réussi.
14. Cliquez sur **Enregistrer**.
Un message confirme la création de la connexion.

Pour afficher un échantillon d'enregistrement extrait à l'aide de la connexion Delimited Flat File, cliquez sur **Prévisualisation** dans la barre d'en-tête. Les enregistrements de fichier sont extraits et les champs triés en fonction des détails que vous avez fournis.

Connexion à Fixed Width Flat File

Pour ajouter une nouvelle connexion Fixed Width Flat File, accédez à **Connexions > Flat File**, puis définissez **Type d'enregistrement** sur **Fixed Width**. Saisissez les détails d'accès du fichier et les détails du type de contenu pour permettre au module Data Federation de le lire correctement.

Remarque : Cette connexion est à utiliser dans le module Metadata Insights.

1. Accédez à **Connexions > Flat File**.
2. Par défaut, l'écran s'ouvre en mode Création. Sinon, cliquez sur  pour ajouter une nouvelle connexion Flat File.
3. Saisissez un **Nom de connexion** pour la connexion aux données Flat File.
4. Saisissez le **Chemin d'accès** en cliquant sur **Parcourir**, puis en sélectionnant le répertoire du fichier.
5. Sélectionnez le **Chiffrement de caractères** du fichier plat dans la liste déroulante.
6. Définissez **Type d'enregistrement** sur **Fixed Width**.
7. Dans le champ **Longueur d'enregistrement**, saisissez le nombre total de caractères d'un enregistrement de fichier.

Répétez l'étape 8 à l'étape 13 pour saisir les détails de tous les champs prévus dans un enregistrement de fichier.

8. Cliquez sur **Ajouter un champ** pour ajouter une ligne pour un champ dans un enregistrement de fichier.
9. Dans le champ **Nom**, saisissez le nom de la valeur de champ.
10. Dans le champ **Type**, sélectionnez le type de données de la valeur de champ.
11. Dans la colonne **Position de début**, saisissez la position de l'enregistrement de fichier à partir de laquelle commence la valeur de champ.

Pour le premier champ d'un enregistrement de fichier, le comptage de la **Position de début** commence à partir de 1.

12. Dans le champ **Longueur**, saisissez le nombre total de caractères couverts par le champ, y compris le caractère à la **Position de début**.

La somme des valeurs **Position de début** et **Longueur** de tout champ doit être inférieure ou égale à la **Longueur d'enregistrement**

Si l'enregistrement de fichier est :

```
01234Rob Smith29PitneyBowes
```

Longueur d'enregistrement = 27

Pour le champ « Nom » :

Position de début = 6

Longueur = 9

```
Name = Rob Smith
```

13. Cochez la case **Raccourcir** si vous souhaitez supprimer tous les espaces blancs au début et/ou à la fin d'une valeur de champ.
14. Cliquez sur **Test**.
Un message confirme que le test de la connexion a réussi.
15. Cliquez sur **Enregistrer**.
Un message confirme la création de la connexion.

Pour afficher un échantillon d'enregistrement extrait à l'aide de la connexion Fixed Width Flat File, cliquez sur **Prévisualisation** dans la barre d'en-tête. Les enregistrements de fichier sont extraits et les champs triés en fonction des détails que vous avez fournis.

Formats de date et d'heure d'une connexion à un fichier

Lors de la lecture des valeurs de date et d'heure des fichiers à l'aide d'une connexion à un fichier dans Spectrum™ Technology Platform, les valeurs doivent respecter certains formats de date et d'heure spécifiques.

Formats de date et d'heure acceptés

- Date : « aaaa-mm-jj »
- Date et heure : « aaaa-mm-jj HH:mm:ss »
- Heure : « HH:mm:ss »

Il s'agit des notations de date et d'heure standard.

Fichiers délimités

Si la fonctionnalité **Détecter le type** est activée lors de la configuration de la connexion Delimited File, les valeurs de date et d'heure des enregistrements de fichier qui respectent les formats susmentionnés sont automatiquement détectées comme de type Date.

Si une valeur date-heure n'est pas conforme à l'un des formats acceptés, la valeur est lue comme une valeur de type Chaîne au lieu d'une valeur de type Date.

Fichiers de largeur fixe

Pour les fichiers de largeur fixe, les valeurs de type Date sont configurées lors de la création de la connexion au fichier de largeur fixe. Par conséquent, ces valeurs sont lues comme des valeurs de type Date, qu'elles respectent ou non les formats acceptés.

Si la valeur date-heure d'un fichier de longueur fixe ne respecte pas les formats acceptés, elle doit être traitée à l'aide de **Transformations** dans le stage Logical Model Creation en appliquant la fonction de catégorie *Conversion* à la valeur :

```
parsedate(String date, String format)
```

Dans ce cas, la *date* est la valeur reçue du fichier, tandis que le *format* est le format de date-heure dans lequel la valeur est reçue du fichier. Cela permet d'analyser correctement la valeur date-heure.

Par exemple, si la *date* = 23-fév-2008, le *format* = jj-*MMM*-aaaa.

Formats de valeur obtenus

Lors de l'aperçu des données dans un Model Store :


- Si la valeur a été lue comme une valeur date/heure, elle apparaît dans l'aperçu sous l'un des formats de date/heure acceptés.
- Si la valeur a été lue comme une valeur Chaîne, elle apparaît comme telle dans l'aperçu.

Connexion à un serveur FTP

Afin que Spectrum™ Technology Platform puisse accéder aux fichiers d'un serveur FTP, vous devez définir une connexion au serveur FTP à l'aide de Management Console. Une fois que cela est fait, vous pouvez créer des flux de données dans Enterprise Designer qui peuvent lire ou écrire des données dans les fichiers sur le serveur FTP.

Avant de vous connecter à un serveur FTP, vérifiez que les paramètres de délai d'attente de votre serveur FTP sont appropriés pour les jobs qui vont utiliser cette connexion. Suivant la conception d'un job, il peut se produire des périodes d'inactivité de la connexion, qui peuvent entraîner une temporisation de la connexion. Par exemple, vous pouvez disposer d'un flux de données avec deux stages Read From File connectés à un stage Import To Hub. Lorsque le stage Import To Hub lit les enregistrements d'un stage Read From File, l'autre est inactif, entraînant éventuellement une temporisation de la connexion au serveur FTP. Pour éviter la temporisation des connexions, vous pouvez envisager de définir la valeur de délai d'attente de votre serveur FTP sur 0.

Remarque : Le serveur FTP doit être exécuté en mode de connexion active. Le mode de connexion passive n'est pas pris en charge.

1. Ouvrez Management Console.
2. Accédez à **Ressources > Sources de données**.
3. Cliquez sur le bouton Ajouter .
4. Dans le champ **Nom**, entrez un nom pour la nouvelle connexion. Ce nom est entièrement de votre choix.


Remarque : Une fois que vous enregistrez une connexion, vous ne pouvez plus modifier le nom.

5. Dans le champ **Type**, sélectionnez **FTP**.
6. Dans les champs **Nom d'utilisateur** et **Mot de passe** saisissez les informations d'identification à utiliser pour authentifier le serveur FTP. Cela n'est nécessaire que si le serveur FTP l'exige.
7. Dans le champ **Hôte**, saisissez le nom d'hôte ou l'adresse IP du serveur FTP.
8. Dans le champ **Port** saisissez le numéro de port réseau utilisé par le serveur FTP.

9. Cliquez sur **Test** pour vérifier que le serveur Spectrum™ Technology Platform peut se connecter au serveur FTP.
10. Cliquez sur **Enregistrer**.

Connexion à Google Cloud Storage

Afin que Spectrum™ Technology Platform puisse accéder aux données de Google Cloud Storage, vous devez définir une connexion à Google Cloud Storage à l'aide de Management Console. Une fois que cela est fait, vous pouvez créer des flux dans Enterprise Designer qui peuvent lire les données de Google Cloud Storage et y écrire des données.

1. Ouvrez Management Console.
2. Accédez à **Ressources > Sources de données**.
3. Cliquez sur le bouton Ajouter .
4. Dans le champ **Nom**, entrez un nom pour la nouvelle connexion. Ce nom est entièrement de votre choix.

Remarque : Une fois que vous enregistrez une connexion, vous ne pouvez plus modifier le nom.

5. Dans le champ **Type**, sélectionnez **Cloud**.
6. Dans le champ **Service Cloud**, sélectionnez **GoogleCloudStorage**.
7. Dans le champ **Nom de compartiment**, saisissez le nom de compartiment tel que défini dans votre service Google Cloud. Il s'agit du compartiment où Spectrum™ Technology Platform lira et écrira des fichiers.
8. Saisissez le nom d'application, le compte de service et le fichier de clé privée fournis par Google.

Remarque : Assurez-vous que le fichier de clé privée est présent sur le serveur Spectrum™ Technology Platform.

9. Vous pouvez définir des droits d'accès dans la section **Droits d'accès**.

Gérer vos données et autorisations Permet à l'utilisateur de gérer les données et les autorisations.

Afficher vos données Permet à l'utilisateur d'afficher les données.

Gérer vos données Permet à l'utilisateur de gérer les données.


10. Pour tester la connexion, cliquez sur **Tester**.
11. Cliquez sur **Enregistrer**.

Pour plus d'informations, reportez-vous à la documentation [Authentification du compte de service](#) de Google.

Connexion à Hadoop

Afin que Spectrum™ Technology Platform puisse accéder aux données de Hadoop, vous devez définir une connexion à Hadoop à l'aide de Management Console. Une fois que cela est fait, vous pouvez créer des flux dans Enterprise Designer qui peuvent lire les données de Hadoop et y écrire des données.

Avertissement : Spectrum™ Technology Platform ne prend pas en charge *Hadoop 2.x* pour Kerberos sur les plates-formes Windows.

1. Ouvrez Management Console.
2. Accédez à **Ressources > Sources de données**.
3. Cliquez sur le bouton Ajouter .
4. Dans le champ **Nom**, entrez un nom pour la nouvelle connexion. Ce nom est entièrement de votre choix.

Remarque : Une fois que vous enregistrez une connexion, vous ne pouvez plus modifier le nom.

5. Dans le champ **Type**, sélectionnez **HDFS**.
6. Dans le champ **Hôte**, saisissez le nom d'hôte ou l'adresse IP de NameNode dans le cluster HDFS.
7. Dans le champ **Port**, entrez le numéro du port réseau.
8. Dans **Utilisateur**, sélectionnez l'une des options suivantes :

Utilisateur serveur	Sélectionnez cette option si l'authentification est activée dans votre cluster HDFS. Cette option utilisera les noms d'utilisateur sous lequel le serveur Spectrum™ Technology Platform est exécuté pour s'authentifier dans HDFS.
Nom utilisateur	Sélectionnez cette option si l'authentification est désactivée dans votre cluster HDFS.
9. Cochez la case **Kerberos** si vous souhaitez activer la fonctionnalité d'authentification Kerberos pour cette connexion au serveur de fichiers HDFS.
10. Si vous avez choisi d'activer l'authentification **Kerberos**, saisissez le chemin d'accès au fichier keytab dans le champ **Chemin d'accès au fichier Keytab**.

Remarque : Assurez-vous que le fichier de clé TAB est présent sur le serveur Spectrum™ Technology Platform.

11. Dans le champ **Protocole**, sélectionnez l'une des options suivantes :

WEBHDFS	Sélectionnez cette option si le cluster HDFS exécute HDFS 1.0 ou version ultérieure. Ce protocole prend en charge les opérations de lecture et d'écriture.
----------------	--

- HFTP** Sélectionnez cette option si le cluster HDFS exécute une version antérieure à HDFS 1.0 ou si votre organisation n'autorise pas le protocole WEBHDFS. Ce protocole ne prend en charge que l'opération de lecture.
- HAR** Sélectionnez cette option pour accéder aux fichiers d'archive Hadoop. Si vous choisissez cette option, indiquez le chemin au fichier d'archive dans le champ **Path**. Ce protocole ne prend en charge que l'opération de lecture.

12. Développez **Options avancées**.

13. Si vous avez sélectionné le protocole WEBHDFS, vous pouvez spécifier les options avancées suivantes, le cas échéant :

Facteur de réplication Indique le nombre de nœuds de données dans lesquels répliquer chaque bloc. Par exemple, le paramètre par défaut de 3 réplique chaque bloc en trois nœuds différents dans le cluster. Le facteur de réplication maximal est de 1 024.

Taille du bloc Indique la taille de chaque bloc. HDFS décompose un fichier en blocs de la taille que vous indiquez ici. Par exemple, si vous indiquez la valeur par défaut 64 Mo, chaque fichier est décomposé en blocs de 64 Mo. Chaque bloc est ensuite répliqué en un nombre de nœuds dans le cluster spécifié dans le champ **Replication factor**.

Permissions de fichier Indique le niveau d'accès aux fichiers écrits dans le cluster HDFS par Spectrum™ Technology Platform. Vous pouvez spécifier des droits de lecture et d'écriture pour chacune des options suivantes :



Remarque : L'autorisation *Exécuter* n'est pas applicable à Spectrum™ Technology Platform.

Utilisateur Il s'agit de l'utilisateur indiqué ci-dessus, **Server user** ou l'utilisateur spécifié dans le champ **User Name**.

Groupe Cela fait référence à un groupe dont l'utilisateur est membre. Par exemple, si l'utilisateur est john123, les permissions Groupe s'appliquent à tout groupe dont john123 est membre.

Autre Cela fait référence à tout autre utilisateur, ainsi qu'aux groupes dont l'utilisateur spécifié n'est pas membre.

14. Dans la grille sous le tableau **Permissions de fichier**, spécifiez les propriétés du serveur pour Hadoop, afin de garantir que les fonctions de tri et de filtrage fonctionnent comme vous le souhaitez lorsque la connexion est utilisée dans un stage ou une activité. Pour ajouter des propriétés, procédez comme suit :

- Cliquez sur  et ajoutez les propriétés et leurs valeurs respectives dans les champs **Propriété** et **Valeur**.
- Cliquez sur  et chargez votre fichier de configuration XML. Le fichier XML doit être semblable à hdfs-site.xml, yarn-site.xml ou core-site.xml.

Remarque : Le fichier de configuration doit être placé sur le serveur.

Ce tableau décrit les propriétés et leurs valeurs, selon le stage ou l'activité qui va utiliser la connexion Hadoop. Les propriétés dépendent également de la version Hadoop utilisée (*Hadoop 1.x* ou *Hadoop 2.x*).

Stage ou activité utilisant la connexion HDFS Propriétés du serveur requis

- Stage **Read from Sequence File**
- Activité **Run Hadoop Pig**

Stage ou activité utilisant la connexion HDFS Propriétés du serveur requis

Paramètres d'Hadoop 1.x

fs.default.name

Indique le nœud et le port d'exécution de Hadoop.

Par exemple,
`hdfs://152.144.226.224:9000`

mapred.job.tracker

Indique le nom d'hôte ou l'adresse IP et le port d'exécution du suiveur de job MapReduce. Si le nom d'hôte est saisi comme local, les jobs sont exécutés sous la forme d'une seule carte et réduisent la tâche.

Par exemple, `152.144.226.224:9001`

dfs.namenode.name.dir

Indique l'endroit où un nœud de nom DFS doit stocker la table nommée dans le système de fichiers local. S'il s'agit d'une liste de répertoires délimitée par des virgules, la table nommée est copiée dans tous les répertoires, à des fins de redondance.

Par exemple,
`file:/home/hduser/Data/namenode`

dfs.datanode.data.dir

Indique l'endroit où un nœud de nom DFS doit stocker ses blocs dans le système de fichiers local. S'il s'agit d'une liste de répertoires délimitée par des virgules, les données sont stockées dans tous les répertoires nommés qui se trouvent généralement sur différents appareils. Les répertoires inexistantes sont ignorés.

Par exemple,
`file:/home/hduser/Data/datanode`

hadoop.tmp.dir

Indique l'emplacement de base des autres répertoires temporaires.

Stage ou activité utilisant la connexion HDFS Propriétés du serveur requis

Par exemple, `/home/hduser/Data/tmp`

Paramètres d'Hadoop 2.x

Stage ou activité utilisant la connexion HDFS Propriétés du serveur requis

fs.defaultFS

Indique le nœud et le port d'exécution de Hadoop.

Par exemple,

```
hdfs://152.144.226.224:9000
```

Remarque : Pour les versions 11.0 et antérieures de Spectrum, le nom de paramètre `fs.defaultfs` doit être utilisé. Notez la différence de casse.

Pour les versions 11 SP1 et suivantes, les deux noms `fs.defaultfs` et `fs.defaultFS` sont valides. Il est recommandé d'utiliser le nom de paramètre `fs.defaultFS` pour Spectrum™ Technology Platform 11 SP1 et les versions suivantes.

yarn.resourcemanager.resource-tracker.address

Indique le nom d'hôte ou l'adresse IP de Resource Manager.

Par exemple, `152.144.226.224:8025`

yarn.resourcemanager.scheduler.address

Indique l'adresse de Scheduler Interface.

Par exemple, `152.144.226.224:8030`

yarn.resourcemanager.address

Indique l'adresse de l'interface Applications Manager contenue dans Resource Manager.

Par exemple, `152.144.226.224:8041`

mapreduce.jobhistory.address

Indique le nom d'hôte ou l'adresse IP et le port d'exécution de MapReduce Job History Server.

Stage ou activité utilisant la connexion HDFS Propriétés du serveur requis

Par exemple, 152.144.226.224:10020

mapreduce.application.classpath

Indique le CLASSPATH des applications MapReduce. Ce CLASSPATH indique l'emplacement des classes associées aux applications MapReduce.

Remarque : Les entrées doivent être séparées par des virgules.

Par exemple

```
$HADOOP_CONF_DIR,
$HADOOP_COMMON_HOME/share/hadoop/common/*,
$HADOOP_COMMON_HOME/share/hadoop/common/lib/*,
$HADOOP_HDFS_HOME/share/hadoop/hdfs/*,
$HADOOP_HDFS_HOME/share/hadoop/hdfs/lib/*,
$HADOOP_MAPRED_HOME/share/hadoop/mapreduce/*,
$HADOOP_MAPRED_HOME/share/hadoop/mapreduce/lib/*,
$HADOOP_YARN_HOME/share/hadoop/yarn/*,
$HADOOP_YARN_HOME/share/hadoop/yarn/lib/*
```

mapreduce.app-submission.cross-platform

Gère plusieurs problèmes de plate-forme qui se produisent si votre serveur Spectrum est exécuté sur un ordinateur Windows et que vous installez Cloudera dessus. Si votre serveur Spectrum et Cloudera sont exécutés sur des systèmes d'exploitation différents, définissez la valeur de ce paramètre sur `true`. Sinon, définissez-le sur `false`.

Remarque : Cloudera ne prend pas en charge les clients Windows. La configuration de ce paramètre contourne le problème et ne constitue pas une solution à tous les problèmes de plate-forme qui se produisent.

Si vous avez coché la case **Kerberos** ci-dessus, ajoutez les ci-dessous les propriétés de configuration Kerberos supplémentaires ci-dessous :

Stage ou activité utilisant la connexion HDFS Propriétés du serveur requis

hadoop.security.authentication

Type de sécurité d'authentification utilisé.
Saisissez la valeur `kerberos`.

yarn.resourcemanager.principal

Principal Kerberos utilisé pour le
Gestionnaire de ressources pour votre
négociateur de ressources Hadoop YARN.

Par exemple, `yarn/_HOST@HADOOP.COM`

dfs.namenode.kerberos.principal

Principal Kerberos utilisé pour le namenode
de votre Hadoop Distributed File System
(HDFS).

Par exemple, `hdfs/_HOST@HADOOP.COM`

dfs.datanode.kerberos.principal

Principal Kerberos utilisé pour le datanode
de votre Hadoop Distributed File System
(HDFS).

Par exemple, `hdfs/_HOST@HADOOP.COM`

Stage ou activité utilisant la connexion HDFS Propriétés du serveur requis

- Stage **Read from File**
- Stage **Write to File**
- Stage **Read from Hive ORC File**
- Stage **Read from Hive ORC File**

Paramètres d'Hadoop 1.x

fs.default.name

Indique le nœud et le port d'exécution de Hadoop.

Par exemple,

`hdfs://152.144.226.224:9000`

Paramètres d'Hadoop 2.x

fs.defaultFS

Indique le nœud et le port d'exécution de Hadoop.

Par exemple,

`hdfs://152.144.226.224:9000`

Remarque : Pour les versions 11.0 et antérieures de Spectrum, le nom de paramètre `fs.defaultfs` doit être utilisé. Notez la différence de casse.

Pour les versions 11 SP1 et suivantes, les deux noms `fs.defaultfs` et `fs.defaultFS` sont valides. Il est recommandé d'utiliser le nom de paramètre `fs.defaultFS` pour Spectrum™ Technology Platform 11 SP1 et les versions suivantes.

15. Pour tester la connexion, cliquez sur **Tester**.

16. Cliquez sur **Enregistrer**.


Après avoir défini une connexion à un cluster HDFS, celle-ci devient disponible dans les stages source et de collecteur de données dans Enterprise Designer, tels que Read from File et Write to File. Vous pouvez sélectionner le cluster HDFS lorsque vous cliquez sur **Ordinateur distant** lors de la définition d'un fichier dans un stage source ou de collecteur de données.

Prise en charge de la compression pour Hadoop

Spectrum™ Technology Platform prend en charge les formats de compression `gzip (.gz)` et `bzip2 (.bz2)` sur Hadoop. Lors de l'utilisation des stages **Read from File** et **Write to File** avec une connexion HDFS, incluez l'extension correspondant au format de compression requis (`.gz` ou `.bz2`) dans le champ **Nom de fichier**. Le fichier est décompressé ou compressé en fonction de l'extension de compression spécifiée. Spectrum™ Technology Platform gère la compression et la décompression des fichiers.

Connexion à une base de données JDBC

Afin que Spectrum™ Technology Platform puisse accéder aux données d'une base de données JDBC, vous devez définir une base de données à l'aide de Management Console. Une fois que cela est fait, vous pouvez créer des flux dans Enterprise Designer qui peuvent lire les données d'une base de données JDBC et y écrire des données.

1. Ouvrez Management Console.
2. Accédez à **Ressources > Sources de données**.
3. Cliquez sur le bouton Ajouter .
4. Dans le champ **Nom**, entrez un nom pour la nouvelle connexion. Ce nom est entièrement de votre choix.

Remarque : Une fois que vous enregistrez une connexion, vous ne pouvez plus modifier le nom.

5. Dans le champ **Type**, sélectionnez le type de base de données auquel vous souhaitez vous connecter.


Le module Spectrum™ Technology Platform Data Integration comprend les pilotes JDBC pour les bases de données SQL Server, Oracle et PostgreSQL. Pour vous connecter à un autre type de base de données, vous devez ajouter le pilote JDBC avant de définir une connexion.

6. Dans le champ **URL**, saisissez l'URL de connexion JDBC. Votre administrateur de base de données peut fournir cette URL.

Par exemple, pour vous connecter à une base de données MySQL nommée « SampleDatabase », hébergée sur un serveur nommé « MyServer », vous devez saisir :

```
jdbc:mysql://MyServer/SampleDatabase
```

7. Il se peut que vous deviez renseigner des champs supplémentaires, suivant le pilote JDBC. Les champs représentent les propriétés de la chaîne de connexion du pilote JDBC que vous avez sélectionné dans le champ **Type**. Reportez-vous à la documentation du fournisseur de pilote JDBC ou contactez votre administrateur de base de données pour avoir plus d'informations sur les propriétés de connexion spécifiques et les valeurs requises par le type de connexion.
8. Cliquez sur **Enregistrer**.

- Testez la connexion en cochant la case en regard de la nouvelle connexion et en cliquant sur le bouton Test .

Importation d'un pilote JDBC

Spectrum™ Technology Platform peut accéder aux données à partir de toute base de données via un pilote JDBC. Les pilotes pour SQL, Oracle et PostgreSQL sont fournis avec le module Spectrum™ Technology Platform Data Integration, ainsi que des pilotes pour d'autres types de bases de données. Si le module Spectrum™ Technology Platform Data Integration n'est pas fourni avec le pilote du type de base de données dont vous avez besoin, vous pouvez ajouter un pilote JDBC.

Dans cette procédure, vous allez importer un pilote JDBC en copiant les fichiers du pilote sur le serveur Spectrum™ Technology Platform. Une fois la procédure terminée, le pilote pourra être utilisé lors de la définition d'une connexion à une base de données JDBC dans Management Console.

Remarque : Cette procédure fonctionne pour les pilotes JDBC 4.x. Si le pilote que vous souhaitez ajouter utilise une version antérieure de JDBC, vous devez ajouter le pilote manuellement dans Management Console. (Pour plus d'informations, reportez-vous à la section [Ajout manuel d'un pilote JDBC](#) à la page 36.)

- Placez tous les fichiers de pilote JDBC pour la base de données dans un dossier nommé :

Name.jdbc

Où *Name* est le nom de votre choix. Le nom du dossier doit se terminer par *.jdbc*.

- Connectez-vous au serveur exécutant Spectrum™ Technology Platform.
- Copiez le dossier contenant le pilote dans le dossier suivant :

Spectrum Location\server\app\drivers

Le pilote est automatiquement importé.

- Pour vérifier que le pilote a été correctement importé, connectez-vous à Management Console et accédez à **Système > Pilotes**. Le pilote doit être répertorié.

Si le pilote n'est pas répertorié, ouvrez le journal système dans Management Console et recherchez les erreurs liées au déploiement de pilotes JDBC.


Ajout manuel d'un pilote JDBC

Spectrum™ Technology Platform peut accéder aux données à partir de toute base de données via un pilote JDBC. Les pilotes pour SQL, Oracle et PostgreSQL sont fournis avec le module Spectrum™ Technology Platform Data Integration, ainsi que des pilotes pour d'autres types de bases de données. Si le module Spectrum™ Technology Platform Data Integration n'est pas fourni avec le pilote du type de base de données dont vous avez besoin, vous pouvez ajouter un pilote JDBC.

Dans cette procédure, vous allez ajouter des fichiers de pilote JDBC sur le serveur, puis définir manuellement la chaîne de connexion et les propriétés de connexion. Avant de commencer, assurez-vous que vous comprenez le format et les propriétés de la chaîne de connexion requis par

le pilote. Pour que le pilote puisse fonctionner, vous devez les définir correctement dans l'ordre. Vous pouvez généralement trouver les informations sur la chaîne de connexion et les propriétés d'un pilote sur le site Web du fournisseur du pilote.

Remarque : Nous vous recommandons d'utiliser cette procédure uniquement lors de l'ajout d'un pilote JDBC qui utilise JDBC 1.x, 2.x ou 3.x. Si le pilote utilise JDBC 4.x, nous vous recommandons d'utiliser la méthode d'importation pour ajouter le pilote. Pour plus d'informations, reportez-vous à la section [Importation d'un pilote JDBC](#) à la page 36.

1. Ouvrez Management Console.
2. Accédez à **Systeme > Pilotes**.
3. Cliquez sur le bouton Ajouter .
4. Dans le champ **Nom**, saisissez un nom pour le pilote. Ce nom est entièrement de votre choix.
5. Dans le champ **JDBC driver class name**, saisissez le nom de la classe Java du pilote. En général, vous pouvez trouver le nom de la classe dans la documentation de votre pilote JDBC.

Par exemple, pour utiliser le pilote JDBC Microsoft, vous pouvez saisir les éléments suivants :

```
com.microsoft.sqlserver.jdbc.SQLServerDriver
```

6. Dans le champ **Modèle de chaîne de connexion**, saisissez l'URL de connexion JDBC à utiliser pour vous connecter à la base de données, y compris toutes les propriétés que vous souhaitez définir dans la chaîne de connexion. Les différents fournisseurs de bases de données utilisent des chaînes de connexion différentes ; c'est pourquoi vous devez consulter la documentation de votre base de données pour obtenir plus d'informations sur la chaîne de connexion.

Si le pilote doit être utilisé par plusieurs connexions de base de données, vous pouvez utiliser des jetons de propriété dans la chaîne de connexion au lieu des valeurs de propriété codées en dur qui peuvent varier pour chaque connexion. Par exemple, si vous souhaitez que certaines connexions utilisent le chiffrement et d'autres non, vous pouvez définir un jeton de propriété pour la propriété de chiffrement.

Pour utiliser un jeton de propriété dans la chaîne de connexion, appliquez la syntaxe suivante :

```
}${PropertyToken}
```

Tout jeton de propriété que vous incluez dans le modèle de chaîne de connexion constituera un champ obligatoire lors de la définition d'une connexion à une base de données.

Remarque : Utilisez le nom de jeton de propriété `${mot de passe}` pour la propriété qui contiendra le mot de passe de base de données. Grâce à ce nom de jeton, le mot de passe est masqué dans le champ dans Management Console et est chiffré dans la base de données.

Par exemple, cette chaîne de connexion SQL contient des jetons de propriété pour l'hôte, le port, l'instance et le chiffrement :

```
jdbc:sqlserver://${host}:${port};databaseName=${instance};encrypt=${encryption};TrustServerCertificate=true
```

Ces jetons constituent des champs obligatoires lors de la définition d'une connexion à une base de données utilisant ce pilote :

[Accueil](#) > [Ressources : Sources de données](#) > [Ajouter une source de don](#)

Ajouter une source de données

*Nom

Connexion


*Type

*Host

*Port

*Instance

*encryption

7. S'il existe des propriétés que vous souhaitez rendre facultatives pour les connexions aux bases de données, définissez-les dans la section **Propriétés de connexion**.
 - a) Dans la section **Propriétés de connexion**, cliquez sur le bouton Ajouter .
 - b) Dans le champ **Libellé**, saisissez une description conviviale de la propriété. Le libellé que vous saisissez ici est utilisé comme libellé de champ dans la fenêtre Connexions lors de la création d'une connexion à l'aide de ce pilote.

- c) Dans le champ **Jeton de propriété**, saisissez le jeton de la propriété facultative. Pour connaître les propriétés prises en charge par le pilote, reportez-vous à la documentation du pilote de base de données.

Remarque : Utilisez le nom de jeton de propriété `mot de passe` pour la propriété qui contiendra le mot de passe de base de données. Grâce à ce nom de jeton, le mot de passe est masqué dans le champ dans Management Console et est chiffré dans la base de données.

Par exemple, si vous souhaitez rendre le chiffrement facultatif pour les connexions aux bases de données utilisant ce pilote, vous pouvez définir la propriété de chiffrement comme suit :

Accueil > Système : pilotes > Modifier un pilote

Modifier le pilote déployé

*Nom

com.mysql.jdbc.Driver.5.1

*Nom de classe du pilote JDBC ?

com.mysql.jdbc.Driver

*Modèle de chaîne de connexion ?

jdbc:mysql://\${host}/\${instance}

Propriétés et pilotes

Propriétés de la connexion ?



Libellé	Code de propriété
<input type="checkbox"/> username	user
<input type="checkbox"/> password	password
<input type="checkbox"/> Use SSL	useSSL

Lorsqu'une connexion à une base de données utilise ce pilote, la propriété de chiffrement est affichée sous forme de propriété facultative dans la connexion à la base de données :

[Accueil](#) > [Ressources : Sources de données](#) > [Ajouter une source de données](#)

Ajouter une source de données

*Nom

MyConnection

Connexion

*Type

com.mysql.jdbc.Driver.5.1


*Host

*Instance

User Name

Password

Use SSL

- Connectez-vous au serveur exécutant Spectrum™ Technology Platform et placez le fichier de pilote de base de données dans un dossier sur le serveur. L'emplacement n'a pas d'importance.
- Dans la section **Fichiers du pilote**, cliquez sur le bouton Ajouter .
- Dans le champ **File path**, saisissez le chemin d'accès au fichier de pilote de base de données sur le serveur.

11. Cliquez sur **Enregistrer**.

Suppression d'un pilote JDBC importé

Il n'est pas possible de supprimer des pilotes JDBC via Management Console si le pilote JDBC a été importé dans Spectrum™ Technology Platform au lieu d'avoir été ajouté manuellement dans Management Console. Au lieu de cela, suivez cette procédure pour supprimer le pilote.

Important : Avant de supprimer un pilote, assurez-vous qu'il n'existe aucune connexion à une base de données via le pilote.

1. Arrêtez le serveur Spectrum™ Technology Platform.
2. Accédez à ce dossier :

```
Spectrum Location\server\app\drivers
```

3. Dans le dossier `drivers`, supprimez le dossier contenant le pilote.
4. Démarrez le serveur Spectrum™ Technology Platform.
5. Pour vérifier que le pilote a été supprimé, connectez-vous à Management Console, accédez à **System > Drivers**, et vérifiez que le pilote n'est plus répertorié.

Types de données de base de données pris en charge

Spectrum™ Technology Platform prend en charge les types de données suivants, couramment utilisés dans les bases de données :

bigdecimal	Type de données numérique prenant en charge 38 points de précision décimaux. Utilisez ce type de données pour les données utilisées dans les calculs mathématiques requérant un haut niveau de précision, en particulier ceux impliquant des données financières. Le type de données <code>bigdecimal</code> prend en charge des calculs plus précis que le type de données <code>double</code> .
booléen	Type logique doté de deux valeurs : <code>True</code> et <code>False</code> .
date	Type de données contenant un mois, un jour et une année. Par exemple, <code>2012-01-30</code> ou <code>January 30, 2012</code> . Vous pouvez indiquer un format de date par défaut dans Management Console.
datetime	Type de données contenant un mois, un jour, une année et des heures, des minutes et des secondes. Par exemple, <code>2012/01/30 6:15 PM</code> .
double	Un type de données numérique qui contient à la fois des nombres en double précision négatifs et positifs entre 2^{-1074} et $(2 \cdot 2^{-52}) \times 2^{1023}$. En notation E, la plage de valeurs est comprise entre <code>-1.79769313486232E+308</code> et <code>1.79769313486232E+308</code> .
flottant	Un type de données numérique contenant des nombres de précision seuls négatifs et positifs entre 2^{-149} et $(2 \cdot 2^{-23}) \times 2^{127}$. En notation E, la plage de valeurs est comprise entre <code>-3.402823E+38</code> et <code>3.402823E+38</code> .
integer	Un type de données numériques qui contient à la fois des nombres entiers négatifs et positifs entre -2^{31} (<code>-2,147,483,648</code>) et $2^{31}-1$ (<code>2,147,483,647</code>).

long	Un type de données numériques qui contient à la fois des nombres entiers négatifs et positifs entre -2^{63} (-9,223,372,036,854,775,808) et $2^{63}-1$ (9,223,372,036,854,775,807).
chaîne	Séquence de caractères.
time	Type de données contenant l'heure du jour. Par exemple, 21:15:59 ou 9:15:59 PM.
Raw	Type de données Oracle pour le stockage de données binaires) longueur variable. La taille maximale est de 2 000 octets (la longueur maximale dans Oracle 7 était de 255 octets).

D'autres types de données de base de données sont automatiquement mappés vers l'un des types de données pris en charge comme suit :

Type de données de base de données	Type de données prises en charge
Types Date/Time	
TIMESTAMP	datetime
Types chaînes	
CHAR	string
CLOB	string
LONGVARCHAR	string
NCHAR	string
NVARCHAR	string
VARCHAR	string
Types numériques	
BIGINT	long
DECIMAL	double
FLOAT	double
NUMERIC	bigdecimal

Type de données de base de données	Type de données prises en charge
REAL	flottant
SMALLINT	integer
TINYINT	integer
Types Booléen	
BIT	booléen

Prise en charge des types de données de base de données pour le module Location Intelligence

Ces types de données de base de données sont automatiquement mappés vers l'un des types de données pris en charge du module Location Intelligence.

Type de données de base de données	Type de données prises en charge
SQL Server	
tinyint	SHORT_INTEGER
smallint	SHORT_INTEGER
int	INTEGER
bigint	LONG_INTEGER
flottant	DOUBLE
real	DOUBLE
decimal(10, 5)	DOUBLE
numeric(10, 5)	DOUBLE
date	DATE
time	TIME

Type de données de base de données	Type de données prises en charge
datetime	DATE_TIME
smalldatetime	DATE_TIME
char (10)	STRING
varchar(10)	STRING
nchar(10)	STRING
nvarchar(10)G	STRING
binary(10)	BINARY
varbinary(10)	BINARY
PostGIS	
smallint	SHORT_INTEGER
integer	INTEGER
bigint	LONG_INTEGER
numeric(10, 5)	DOUBLE
real	DOUBLE
double precision	DOUBLE
serial	INTEGER
bigserial	LONG_INTEGER
bytea	BINARY
date	DATE

Type de données de base de données	Type de données prises en charge
time	TIME
timestamp	DATE_TIME
character(10)	STRING
character varying(10)	STRING
nchar(10)	STRING
Oracle	
NUMBER	DOUBLE
CHAR(10)	STRING
VARCHAR(10)	STRING
VARCHAR2(10)	STRING
NCHAR(10)	STRING
NVARCHAR2(10)	STRING
DATE	DATE_TIME
TIMESTAMP	DATE_TIME
BLOB	BINARY
SAP HANA	
tinyint	SHORT_INTEGER
smallint	SHORT_INTEGER


Type de données de base de données	Type de données prises en charge
integer	INTEGER
bigint	LONG_INTEGER
smalldecimal	DOUBLE
decimal(10, 5)	DOUBLE
real	DOUBLE
double	DOUBLE
float(30)	DOUBLE
varchar(30)	STRING
nchar(10)	STRING
nvarchar(30)	STRING
alphanum(30)	STRING
date	DATE
time	TIME
seconddate	DATE_TIM
timestamp	DATE_TIM
varbinary(30)	BINARY

Restrictions

- Les connecteurs MongoDB/Cassandra ne sont pas pris en charge via PrestoDB dans Metadata Insights. Il existe des connecteurs distincts pour MongoDB et Cassandra.
- Write to Any DB via Presto n'est pas recommandé par Presto DB et n'est donc pas pris en charge par le connecteur Presto JDBC.

Connexion à Knox

Une passerelle Apache Knox vous permet d'accéder à un service Hadoop via une couche de sécurité Knox. Afin que Spectrum™ Technology Platform puisse accéder aux données de Hadoop via Knox, vous devez définir une connexion à Knox à l'aide de Management Console. Une fois que cela est fait, vous pouvez créer des flux dans Enterprise Designer qui peuvent lire les données de Hadoop et y écrire des données via Knox.

1. Ouvrez Management Console.
2. Accédez à **Ressources** > **Sources de données**.
3. Cliquez sur le bouton Ajouter .
4. Dans le champ **Nom**, entrez un nom pour la nouvelle connexion. Ce nom est entièrement de votre choix.

Remarque : Une fois que vous enregistrez une connexion, vous ne pouvez plus modifier le nom.

5. Dans le champ **Type**, sélectionnez *Passerelle*.
6. Dans le champ **Type de passerelle**, sélectionnez **Knox**.
7. Dans le champ **Hôte**, saisissez le nom d'hôte ou l'adresse IP du nœud du cluster HDFS exécutant la passerelle.
8. Dans le champ **Port**, saisissez le numéro de port de la passerelle Knox.
9. Dans le champ **Nom d'utilisateur**, saisissez le nom d'utilisateur de la passerelle Knox.
10. Dans le champ **Mot de passe**, saisissez le mot de passe vous autorisant à accéder à la passerelle Knox.
11. Dans le champ **Nom de passerelle**, saisissez le nom de la passerelle Knox à laquelle vous souhaitez accéder.
12. Dans le champ **Nom de cluster**, saisissez le nom du cluster Hadoop auquel accéder.
13. Dans le champ **Protocole**, sélectionnez *webhdfs*.
14. Dans le champ **Nom de service**, saisissez le nom du service Hadoop auquel accéder.
15. Pour tester la connexion, cliquez sur **Tester**.
16. Cliquez sur **Enregistrer**.


Après avoir défini une connexion Knox à un cluster HDFS, vous pouvez utiliser la connexion dans Enterprise Designer, dans les stages **Read from File** et **Write to File**. Vous pouvez sélectionner le cluster HDFS lorsque vous cliquez sur **Ordinateur distant** lors de la définition d'un fichier dans un stage source ou de collecteur de données.

Connexion à Marketo

Afin que Spectrum™ Technology Platform puisse accéder aux données de Marketo, vous devez définir une connexion à Marketo à l'aide de Management Console. Une fois que cela est fait, vous

pouvez créer des flux dans Enterprise Designer qui peuvent lire les données de Marketo et y écrire des données.

Remarque : Cette connexion est à utiliser dans le module Metadata Insights.

1. Ouvrez Management Console.
2. Accédez à **Ressources > Sources de données**.
3. Cliquez sur le bouton Ajouter .
4. Dans le champ **Nom**, entrez un nom pour la nouvelle connexion. Ce nom est entièrement de votre choix.

Remarque : Une fois que vous enregistrez une connexion, vous ne pouvez plus modifier le nom.

5. Dans le champ **Type**, sélectionnez **Marketo**.
6. Dans le champ **URL d'accès**, saisissez l'URL de votre compte Marketo.

Pour trouver votre URL d'accès, connectez-vous à votre compte Marketo et accédez à **Admin > Intégration > Services Web**. L'URL d'accès se trouve sous l'en-tête **API REST** et se présente au format suivant :

```
https://AccountID.mktorest.com/rest
```

Copiez la partie de l'URL avant /rest. Il peut s'agir par exemple de
https://AccountID.mktorest.com.

7. Saisissez l'ID client et la clé secrète client de votre compte Marketo.

Pour trouver votre ID client et votre clé secrète, connectez-vous à votre compte Marketo et accédez à **Admin > Intégration > LaunchPoint > API Rest > Afficher les détails**. La fenêtre contextuelle affiche les détails.

8. Pour tester la connexion, cliquez sur **Tester**.
9. Cliquez sur **Enregistrer**.

Restrictions de Marketo

1. Cette requête s'applique uniquement aux entités `List` et `Activity_type`. Pour les autres, fournissez le type de filtre.

```
Select * from Marketo_Table
```


2. Marketo ne prend pas en charge les jointures, sauf entre les entités `Lead` et `Lead_List`. La requête de jointure entre `Lead` et `Lead_List` d'un élément `List_Id` se présente comme suit :

```
Select Lead.* from Lead Inner Join Lead_List  
On Lead.ID = Lead_List.Lead_ID  
And Lead_List.List_ID = <List ID>
```


Connexion à Microsoft Dynamics 365

Connexion à Microsoft Dynamics 365 Online

Afin que Spectrum™ Technology Platform puisse accéder aux données de Microsoft Dynamics 365 Online, vous devez définir une connexion à Microsoft Dynamics 365 Online à l'aide de Management Console. Une fois que cela est fait, vous pouvez créer des flux dans Enterprise Designer qui peuvent lire les données de Microsoft Dynamics 365 Online et y écrire des données.

Remarque : Cette connexion est à utiliser dans le module Metadata Insights.

Suivez ces étapes pour définir une connexion à **Microsoft Dynamics 365 Online**.

1. Ouvrez Management Console.
2. Accédez à **Ressources > Sources de données**.
3. Cliquez sur le bouton Ajouter .
4. Dans le champ **Nom**, entrez un nom pour la nouvelle connexion. Ce nom est entièrement de votre choix.

Remarque : Une fois que vous enregistrez une connexion, vous ne pouvez plus modifier le nom.

5. Dans le champ **Type**, sélectionnez **Microsoft Dynamics 365**.
6. Dans le champ **Type de développement**, sélectionnez **En ligne**.
7. Dans le champ **Nom d'utilisateur**, saisissez votre nom d'utilisateur Microsoft Dynamics.
8. Dans le champ **Mot de passe**, saisissez votre mot de passe Microsoft Dynamics.
9. Dans le champ **Nom d'organisation** saisissez votre nom unique d'organisation, qui identifie votre instance CRM.

Pour trouver votre nom unique d'organisation, connectez-vous à Microsoft Dynamics et accédez à **Paramètres > Personnalisation > Personnalisations > Ressources destinées aux développeurs**. Votre nom unique d'organisation s'affiche.

10. Dans le champ **Région** sélectionnez la région géographique de votre compte Microsoft Dynamics.
11. Pour tester la connexion, cliquez sur **Tester**.
12. Cliquez sur **Enregistrer**.

Connexion à Microsoft Dynamics 365 On Premises

Actuellement, Spectrum prend en charge les demandes basées sur les authentifications pour Microsoft Dynamics 365 On Premises.

Pré-requis :

Importer le certificat dans le fichier de magasin de clés : pour importer les certificats de serveur Dynamics CRM dans le magasin de clés de distribution Spectrum Java, procédez comme suit :

1. Copiez les certificats de serveur dans un dossier local.
2. Accédez au chemin d'accès à la distribution Spectrum JAVA suivant : <SPECTRUM_HOME>\java\jre\lib\security.
3. Exécutez la commande suivante pour importer les certificats : <codeph>keytool -importcert -alias <nom


d'alias du certificat> -file " <chemin d'accès au certificat>\<nom du certificat>" -keystore keystore.jks</codeph> dans Windows et <codeph>keytool -import -alias <nom d'alias du certificat> -file "<chemin d'accès au certificat>/<nom du certificat>" -keystore keystore.jks</codeph> dans Unix

Définition d'une connexion à Microsoft Dynamics 365 On Premises

Afin que Spectrum™ Technology Platform puisse accéder aux données de Microsoft Dynamics 365 On Premise, définissez une connexion à Microsoft Dynamics 365 On Premise à l'aide de Management Console. Une fois la connexion définie, vous pouvez créer des flux dans Enterprise Designer qui peuvent lire les données de Microsoft Dynamics 365 On Premise et y écrire des données.

Remarque : Cette connexion est à utiliser dans le module Metadata Insights.

Procédez comme suit pour définir une connexion à **Microsoft Dynamics 365 On Premises** :

1. Ouvrez Management Console.
2. Accédez à **Ressources > Sources de données**.
3. Cliquez sur le bouton Ajouter .
4. Dans le champ **Nom**, entrez un nom pour la nouvelle connexion. Ce nom est entièrement de votre choix.

Remarque : Une fois que vous enregistrez une connexion, vous ne pouvez plus modifier le nom.

5. Cliquez sur **Microsoft Dynamics 365** dans le **Type**.
6. Cliquez sur **On Premise** dans le **Type de développement**.
7. Saisissez votre nom d'utilisateur Microsoft Dynamics dans le **Nom d'utilisateur**.
8. Saisissez votre mot de passe Microsoft Dynamics dans le **Mot de passe**.
9. Saisissez le nom de l'hôte dans le **Nom d'hôte**.

10. Saisissez le nom de port dans le **Nom de port**.
11. Saisissez l'URL de STS dans **URL STS**.
12. Cliquez sur **Test** pour tester la connexion.
13. Cliquez sur **Enregistrer**.


Restrictions

Voici les restrictions :

1. **Créer/Mettre à jour** :Créer/Mettre à jour peut échouer si une colonne d'une entité est mappée vers plusieurs entités de référence. Par exemple, « ParentCustomerId » de Customer peut être associé à Account, Lead, etc. Pour résoudre ce problème, les données de cette colonne doivent se présenter au format suivant : « ReferenceEntityName:GUID » au lieu de « GUID ».

Connexion à un Model Store

Connectez-vous à un Model Store pour utiliser les données fédérées de différentes sources telles que des bases de données, des serveurs de fichiers et des services Cloud. Une fois qu'une connexion est définie, vous pouvez utiliser les données des Logical Models et des Physical Models d'un Model Store dans les stades **Read from DB** et **Write to DB** d'Enterprise Designer.

1. Ouvrez Management Console.
2. Accédez à **Ressources > Sources de données**.
3. Cliquez sur le bouton Ajouter .
4. Dans le champ **Nom**, entrez un nom pour la nouvelle connexion. Ce nom est entièrement de votre choix.

Remarque : Une fois que vous enregistrez une connexion, vous ne pouvez plus modifier le nom.

5. Dans le champ **Type**, sélectionnez **Model Store**.
6. Dans le champ **Model Store**, saisissez le nom du Model Store auquel vous souhaitez vous connecter.

Pour trouver les noms des Model Stores disponibles, ouvrez Metadata Insights sur Modeling, puis cliquez sur l'onglet **Model Store**.

7. Pour tester la connexion, cliquez sur **Tester**.
8. Cliquez sur **Enregistrer**.

Connexion à NetSuite

Afin que Spectrum™ Technology Platform puisse accéder aux données de NetSuite, vous devez définir une connexion à NetSuite à l'aide de Management Console. Une fois que cela est fait, vous pouvez créer des flux dans Enterprise Designer qui peuvent lire les données de NetSuite et y écrire


des données. Lors de la lecture et de l'écriture d'une connexion NetSuite, les deux modes, interactif et par lots, sont pris en charge.

Remarque : Cette connexion est à utiliser dans le module Metadata Insights.

Spectrum™ Technology Platform prend en charge les types d'entité NetSuite suivants :

- Enregistrements standard
- Enregistrements personnalisés
- Recherches enregistrées
- Jointures entre des enregistrements standard

Pour vous connecter à NetSuite :

1. Ouvrez Management Console.
2. Accédez à **Ressources > Sources de données**.
3. Cliquez sur le bouton Ajouter .
4. Dans le champ **Nom**, entrez un nom pour la nouvelle connexion. Ce nom est entièrement de votre choix.

Remarque : Une fois que vous enregistrez une connexion, vous ne pouvez plus modifier le nom.

5. Dans le champ **Type**, sélectionnez **NetSuite**.
6. Dans le champ **Email**, saisissez l'adresse électronique associée au compte NetSuite à utiliser pour la connexion.
7. Dans le champ **Mot de passe**, saisissez le mot de passe du compte NetSuite.
8. Dans le champ **Compte**, saisissez le nom d'utilisateur du compte NetSuite.
9. Dans le champ **Rôle**, sélectionnez le rôle approprié pour cette connexion parmi les rôles mappés vers le compte d'utilisateur NetSuite donné.

Le champ **Rôle** est facultatif. Si vous laissez le champ **Rôle** vide, le rôle par défaut est utilisé pour se connecter via la connexion.

Avertissement : Seuls les rôles standard sont pris en charge. Les rôles personnalisés ne sont pas pris en charge.

10. Pour tester la connexion, cliquez sur **Tester**.
11. Cliquez sur **Enregistrer**.

Remarque : Pour effectuer une opération `INSERT` sur un enregistrement à l'aide d'une connexion NetSuite, utilisez une requête `UPSERT` avec la clé primaire (`internalId`) vide.

Restrictions de NetSuite

1. Lors de l'interrogation à l'aide de jointures, vous devez indiquer des colonnes spécifiques. Par exemple, la requête suivante n'est pas prise en charge :

```
select * from CUSTOMER_M
```

2. Les connexions simultanées à NetSuite ne sont pas prises en charge, car NetSuite autorise une seule connexion par compte.
3. Vous pouvez écrire uniquement des enregistrements standard et personnalisés.
4. Pour les deux requêtes `UPDATE` et `UPSERT`, une opération `UPSERT` est réalisée.
5. Dans le stage Write to DB, la taille de lot maximale autorisée pour une opération `insert` est de 200 et la taille de lot maximale pour une opération `update` est de 100.
- 6.


Connexion à NoSQL

Afin que Spectrum™ Technology Platform puisse accéder aux données d'une base de données NoSQL, vous devez définir une connexion à la base de données NoSQL à l'aide de Management Console.

Les types de base de données NoSQL pris en charge sont les suivants :

1. Couchbase
2. MongoDB

Une fois la connexion NoSQL souhaitée définie, vous pouvez créer des flux dans Enterprise Designer qui peuvent lire les données de cette base de données et y écrire des données.

1. Ouvrez Management Console.
2. Accédez à **Ressources > Sources de données**.
3. Cliquez sur le bouton Ajouter .
4. Dans le champ **Nom**, entrez un nom pour la nouvelle connexion. Ce nom est entièrement de votre choix.


Remarque : Une fois que vous enregistrez une connexion, vous ne pouvez plus modifier le nom.

5. Dans le champ **Type**, sélectionnez l'une des options suivantes :
 - Couchbase
 - MongoDB
6. Spécifiez l'**Hôte**, le **Port**, la **Base de données**, le **Nom d'utilisateur** et le **Mot de passe** de la base de données NoSQL spécifique à laquelle vous souhaitez accéder.
7. Cliquez sur **Test** pour vérifier que la connexion à la base de données fonctionne.
8. Cliquez sur **OK**.

Connexion à Salesforce

Afin que Spectrum™ Technology Platform puisse accéder aux données de Salesforce, vous devez définir une connexion à Salesforce à l'aide de Management Console. Une fois que cela est fait, vous pouvez créer des flux dans Enterprise Designer qui peuvent lire les données de Salesforce et y écrire des données.

Remarque : Cette connexion est à utiliser dans le module Metadata Insights.

1. Ouvrez Management Console.
2. Accédez à **Ressources** > **Sources de données**.
3. Cliquez sur le bouton Ajouter .
4. Dans le champ **Nom**, entrez un nom pour la nouvelle connexion. Ce nom est entièrement de votre choix.

Remarque : Une fois que vous enregistrez une connexion, vous ne pouvez plus modifier le nom.

5. Dans le champ **Type**, sélectionnez **Salesforce**.
6. Dans le champ **Nom d'utilisateur**, saisissez l'ID de courrier électronique enregistré dans le magasin de données Salesforce.
7. Dans le champ **Mot de passe**, saisissez une combinaison du mot de passe du portail Salesforce et du jeton de sécurité généré via le portail Salesforce.

Par exemple, si votre mot de passe est Sales@Test et si le jeton de sécurité fourni par Salesforce est 56709367, le mot de passe d'authentification de cette connexion Salesforce est Sales@Test56709367.

8. Pour tester la connexion, cliquez sur **Tester**.
9. Cliquez sur **Enregistrer**.

Remarque : Les champs d'audit sont activés sur toutes les tables par défaut. Les champs d'audit Salesforce sont les suivants :

- Date de création
- Date de la dernière modification
- Créé par
- Dernière modification effectuée par

Avertissement : Les Physical Models créés dans Spectrum™ Technology Platform, versions 10 et antérieures, à l'aide de connexions Salesforce, doivent être ouvertes et enregistrées de nouveau pour que les champs d'audit de leurs tables soient activés.


Restriction Salesforce

Les fonctions d'agrégation ne sont pas prises en charge lors de l'exécution des requêtes sur un Model Store.

Connexion à SAP NetWeaver

La création d'une connexion SAP NetWeaver dans Management Console via OData Services vous permet de lire, d'écrire et de synchroniser vos données CRM et ERP. Lors de la lecture et de l'écriture d'une connexion SAP, les deux modes, interactif et par lots, sont pris en charge.

Remarque : Cette connexion est à utiliser dans le module Metadata Insights.

1. Ouvrez Management Console.
2. Accédez à **Ressources > Sources de données**.
3. Cliquez sur le bouton Ajouter .
4. Dans le champ **Nom**, entrez un nom pour la nouvelle connexion. Ce nom est entièrement de votre choix.

Remarque : Une fois que vous enregistrez une connexion, vous ne pouvez plus modifier le nom.

5. Dans le champ **Type**, sélectionnez **SAP**.
6. Dans le champ **Nom d'utilisateur**, saisissez le nom d'utilisateur permettant d'accéder au service Web SAP.
7. Dans le champ **Mot de passe**, saisissez le mot de passe du service Web SAP.
8. Dans le champ OdataURL, saisissez l'adresse du service Web Odata à utiliser pour cette connexion.
9. Cliquez sur **Test**.
Un message confirme que le test de la connexion a réussi.
10. Cliquez sur **Enregistrer**.
Un message confirme la création de la connexion.

Remarque : Pour effectuer des opérations d'extraction, un service OData doit prendre en charge les opérations \$skip et \$top. Si le service ne prend pas en charge ces opérations, les enregistrements extraits affichent des incohérences dans l'aperçu Model Store.


Restrictions de SAP NetWeaver

Pour les deux opérations UPDATE et UPSERT, une opération UPDATE est réalisée.

Connexion à SharePoint

Afin que Spectrum™ Technology Platform puisse accéder aux données de SharePoint, vous devez définir une connexion à SharePoint à l'aide de Management Console. Une fois que cela est fait,

vous pouvez créer des flux dans Enterprise Designer qui peuvent lire les données de SharePoint et y écrire des données.

1. Ouvrez Management Console.
2. Accédez à **Ressources > Sources de données**.
3. Cliquez sur le bouton Ajouter .
4. Dans le champ **Nom**, entrez un nom pour la nouvelle connexion. Ce nom est entièrement de votre choix.

Remarque : Une fois que vous enregistrez une connexion, vous ne pouvez plus modifier le nom.

5. Dans le champ **Type**, sélectionnez **Cloud**.
6. Dans le champ **Service Cloud**, sélectionnez **Sharepoint**.
7. Dans le champ **Versión**, sélectionnez **v2010**. Spectrum™ Technology Platform prend actuellement en charge SharePoint version 2010.
8. Dans le champ **Protocole**, sélectionnez le protocole requis pour se connecter Sharepoint.
9. Dans le champ **Adresse du serveur**, saisissez le nom d'hôte ou l'adresse IP du serveur SharePoint auquel vous souhaitez vous connecter.
10. Saisissez le nom d'utilisateur et le mot de passe à utiliser pour l'authentification dans SharePoint.
11. Dans le champ **Projet**, saisissez le projet spécifique correspondant à l'emplacement SharePoint auquel vous souhaitez accéder.
12. Pour tester la connexion, cliquez sur **Tester**.
13. Cliquez sur **Enregistrer**.

Exemple

Par exemple, imaginons que vous souhaitiez créer une connexion à cette URL SharePoint :

```
https://sharepoint.example.com/sites/myportal
```

Vous devez renseigner les champs **Protocole**, **Adresse du serveur** et **Projet** comme suit :


- **Protocole** : https
- **Adresse du serveur** : sharepoint.example.com
- **Projet** : myportal

Connexion à Siebel

Afin que Spectrum™ Technology Platform puisse accéder aux données de Siebel, vous devez définir une connexion à Siebel à l'aide de Management Console. Une fois que cela est fait, vous pouvez

créer des flux dans Enterprise Designer qui peuvent lire les données de Siebel et y écrire des données.

Remarque : Cette connexion est à utiliser dans le module Metadata Insights.

1. Ouvrez Management Console.
2. Accédez à **Ressources > Sources de données**.
3. Cliquez sur le bouton Ajouter .
4. Dans le champ **Nom**, entrez un nom pour la nouvelle connexion. Ce nom est entièrement de votre choix.

Remarque : Une fois que vous enregistrez une connexion, vous ne pouvez plus modifier le nom.

5. Dans le champ **Type**, sélectionnez **Siebel**.
6. Dans le champ **Hôte**, saisissez le nom d'hôte du serveur sur lequel Siebel est installé.
7. Dans le champ **Port**, saisissez le numéro de port d'écoute du composant SCBroker.

Remarque : Pour SCBPort, utilisez 2321.

8. Dans le champ **Serveur**, saisissez le nom de votre serveur Siebel Enterprise.
9. Dans le champ **Gestionnaire d'objets**, saisissez le nom du Gestionnaire d'objets.
10. Saisissez le nom d'utilisateur et le mot de passe à utiliser pour l'authentification dans Siebel.
11. Dans le champ **Paramètres régionaux**, sélectionnez le type de Gestionnaire d'objets et le pack linguistique de l'application Siebel Business.

Remarque : Pour les paramètres régionaux autres que **anglais**, vous devez installer les fichiers JAR de paramètres régionaux spécifiques. Pour plus d'informations, reportez-vous à la section [Fichiers JAR requis](#).

12. Pour tester la connexion, cliquez sur **Tester**.
13. Cliquez sur **Enregistrer**.

Dans un Physical Model, les composants commerciaux Siebel sont affichés au format `<Business Object>.<Business Component>`.

Remarque : Les Business Components qui ne font partie d'aucun Business Object ne sont pas affichés dans le Physical Model.

Restrictions de Siebel

1. Les champs des composants commerciaux Siebel portant le même nom, mais avec une casse différente, sont traités comme suit : la première occurrence reste inchangée, tandis que `_PB_` et une valeur numérique croissante sont ajoutés à l'ensemble des autres occurrences.

Par exemple, les champs tels que `DeDup Token`, `Dedup Token` et `DEdup Token` d'un composant commercial Siebel sont renommés `DeDup Token`, `Dedup Token_PB_1` et `DEdup Token_PB_2`, respectivement.

2. Si un nom de champ contient un point (.) dans un schéma Siebel, le point est supprimé du nom de champ à des fins d'utilisation dans Spectrum™ Technology Platform.
3. Lors de la création d'un Physical Model depuis une connexion Siebel, les liens entre deux composants commerciaux Siebel sont affichés uniquement si le nom de lien correspond au nom de composant commercial du schéma Siebel.


Par exemple, dans un schéma Siebel, si un lien est défini entre les composants commerciaux `Account` et `Contact` alors qu'un autre composant commercial `Contact Custom` utilise les liens `Custom`, dans le Physical Model créé pour cette connexion Siebel, aucun lien n'est affiché entre `Account` et `Contact Custom`. Les liens sont affichés uniquement entre `Account` et `Contact`, car le lien initial figure entre les deux dans le schéma Siebel.

4. Les champs d'un composant commercial indiqués comme inactifs dans le schéma Siebel ne sont pas affichés dans les Physical Models, les Logical Models ni les Model Stores.

Connexion à Splunk

Afin que Spectrum™ Technology Platform puisse accéder aux données de Splunk, vous devez définir une connexion à Splunk à l'aide de Management Console. Une fois que cela est fait, vous pouvez créer des flux dans Enterprise Designer qui peuvent lire les données de Splunk et y écrire des données.

Remarque : Cette connexion est à utiliser dans le module Metadata Insights.

1. Ouvrez Management Console.
2. Accédez à **Ressources > Sources de données**.
3. Cliquez sur le bouton Ajouter .
4. Dans le champ **Nom**, entrez un nom pour la nouvelle connexion. Ce nom est entièrement de votre choix.

Remarque : Une fois que vous enregistrez une connexion, vous ne pouvez plus modifier le nom.

5. Dans le champ **Type**, sélectionnez **Splunk**.
6. Dans le champ **Nom d'utilisateur**, saisissez le nom d'utilisateur du compte Splunk d'authentification de l'instance Splunk.
7. Dans le champ **Mot de passe**, saisissez le mot de passe du compte Splunk.
8. Dans le champ **Hôte**, saisissez l'adresse ou le nom d'hôte du serveur d'hébergement de la source de données Splunk.
9. Dans le champ **Port**, saisissez le numéro de port de la source de données Splunk.
10. Pour tester la connexion, cliquez sur **Tester**.

11. Cliquez sur **Enregistrer**.

Restrictions de Splunk


Cette requête n'est pas prise en charge :

```
select count(*) from SplunkTable
```

Connexion à SuccessFactors

Afin que Spectrum™ Technology Platform puisse accéder aux données de SuccessFactors, vous devez définir une connexion à SuccessFactors à l'aide de Management Console. Une fois que cela est fait, vous pouvez créer des flux dans Enterprise Designer qui peuvent lire les données de SuccessFactors et y écrire des données.

Remarque : Cette connexion est à utiliser dans le module Metadata Insights.

1. Ouvrez Management Console.
2. Accédez à **Ressources > Sources de données**.
3. Cliquez sur le bouton Ajouter .
4. Dans le champ **Nom**, entrez un nom pour la nouvelle connexion. Ce nom est entièrement de votre choix.

Remarque : Une fois que vous enregistrez une connexion, vous ne pouvez plus modifier le nom.

5. Dans le champ **Type**, sélectionnez **SuccessFactors**.
6. Dans le **ID de société**, saisissez votre ID de société, qui identifie l'instance unique de votre société dans un centre de données SuccessFactors spécifique.
7. Dans le champ **URL du service**, saisissez l'URL du serveur SuccessFactors auquel vous souhaitez vous connecter. Cette URL est spécifique au centre de données international auquel l'ID de société est mappé.
8. Saisissez le nom d'utilisateur et le mot de passe de votre instance client SuccessFactors.
9. Pour tester la connexion, cliquez sur **Tester**.
10. Cliquez sur **Enregistrer**.


Restrictions de SuccessFactors

1. Les opérations de traitement par lots peuvent être effectuées uniquement à l'aide des requêtes `upsert`. Par conséquent, les requêtes `insert` et `update` sont également effectuées sous forme de requêtes `upsert` dans les opérations de traitement par lots.
2. Il se peut que les propriétés de la colonne/table, telles qu'indiquées dans le schéma Physical Model d'une connexion SuccessFactors, ne fonctionnent pas comme prévu lors de l'utilisation de l'opération correspondante. Par exemple, une colonne qui a été marquée comme pouvant être mise à jour, peut lever une exception lorsque vous essayez de mettre à jour cette colonne.

Connexion à SugarCRM

Afin que Spectrum™ Technology Platform puisse accéder aux données de SugarCRM, vous devez définir une connexion à SugarCRM à l'aide de Management Console. Une fois que cela est fait, vous pouvez créer des flux dans Enterprise Designer qui peuvent lire les données de SugarCRM et y écrire des données. Les versions de SugarCRM en ligne et locale sont prises en charge.

Remarque : Cette connexion est à utiliser dans le module Metadata Insights.

1. Ouvrez Management Console.
2. Accédez à **Ressources > Sources de données**.
3. Cliquez sur le bouton Ajouter .
4. Dans le champ **Nom**, entrez un nom pour la nouvelle connexion. Ce nom est entièrement de votre choix.

Remarque : Une fois que vous enregistrez une connexion, vous ne pouvez plus modifier le nom.

5. Dans le champ **Type**, sélectionnez **SugarCRM**.
6. Saisissez vos nom d'utilisateur et votre mot de passe SugarCRM.
7. Dans le champ **URL**, saisissez l'URL du compte SugarCRM à utiliser pour cette connexion.
8. Saisissez l'**Id Client** et le **Secret Client** de votre compte SugarCRM.
9. Pour tester la connexion, cliquez sur **Tester**.
10. Cliquez sur **Enregistrer**.

Restrictions de SugarCRM


1. Pour les deux requêtes `UPDATE` et `UPSERT`, une opération `UPSERT` est réalisée.
2. Il est possible que les colonnes **Annulable** et **Actualisable** des propriétés de la table, telles qu'affichées dans le **Schéma Physical Model** de la connexion, ne représentent pas le fonctionnement correct. Par exemple, une colonne non marquée comme actualisable peut lever une exception système lorsque vous essayez de la mettre à jour et, inversement, une colonne marquée comme annulable peut ne pas lever d'exception lors de sa mise à jour.
3. Lors d'une requête utilisant des jointures, vous devez utiliser un alias.

Connexion à Oracle Eloqua

Pour que Spectrum™ Technology Platform puisse accéder aux données dans Oracle Eloqua, vous devez définir une connexion à Eloqua Oracle à l'aide de Management Console.

. Une fois que cela est fait, vous pouvez créer des flux dans Enterprise Designer qui peuvent lire les données d'Eloqua et y écrire des données.

Remarque : Cette connexion est à utiliser dans le module Metadata Insights.

1. Ouvrez Management Console.
2. Accédez à **Ressources > Sources de données**.
3. Cliquez sur le bouton Ajouter .
4. Dans le champ **Nom**, entrez un nom pour la nouvelle connexion. Ce nom est entièrement de votre choix.

Remarque : Une fois que vous enregistrez une connexion, vous ne pouvez plus modifier le nom.

5. Dans le champ **Type**, cliquez sur **Eloqua**.
6. Saisissez le **Nom du site** qui est identique au nom de la société.
7. Saisissez le nom d'utilisateur dans le champ **Nom d'utilisateur**.
8. Saisissez le mot de passe dans le champ **Mot de passe**.
9. Cliquez sur **Test** pour tester la connexion.
10. Cliquez sur **Enregistrer**.

Prise en charge des entités

Le tableau suivant décrit les entités et les opérations prises en charge pour Oracle Eloqua.

Tableau 1 :

Nom de l'entité	Créer	Lire	Mettre à jour	Supprimer	Prise en charge par lots	Taille de lot maximale
Comptes	X	X	X	X	Insérer/Mettre à jour*	1000
Groupes de comptes		X				
Campagne		X				
Contacts	X	X	X	X	Insérer/Mettre à jour*	1000
Liste de contacts	X	X	X	X		
Segment de contacts	X	X	X	X		
E-mails		X				

Nom de l'entité	Créer	Lire	Mettre à jour	Supprimer	Prise en charge par lots	Taille de lot maximale
Dossiers d'e-mails		X				
Groupes d'e-mails		X				
Microsites		X				
Utilisateurs		X				
Visiteurs		X				
Activité						
Ouverture d'e-mails		X				
EmailClickthrough		X				
Envoi d'e-mails		X				
S'abonner		X				
Se désinscrire		X				
Rebond		X				
WebVisit		X				
PageView		X				
FormSubmit		X				
Entités personnalisées						
Contact List Members	X	X		X	Insérer/Supprimer	1000

Nom de l'entité	Créer	Lire	Mettre à jour	Supprimer	Prise en charge par lots	Taille de lot maximale
ContactSegmentMembers		X				

L'opération Mettre à jour fonctionne comme Upsert.

Opérations spéciales

- Utilisez la requête de jointure suivante pour récupérer les contacts d'une Liste de contacts :

```
select * from Contacts inner join ContactListMembers on
Contacts.Eloqua_Contact_ID = ContactListMembers.Contact_Id where
ContactListMembers.ContactList_Id = '<id>'
```

Utilisez la requête de jointure suivante pour récupérer les contacts d'un Segment de contacts :

```
select * from Contacts inner join ContactSegmentMembers on
Contacts.Eloqua_Contact_ID = ContactSegmentMembers.Contact_Id where
ContactSegmentMembers.Contactlist_Id = '<id>'
```

- Utilisez l'instruction suivante pour insérer des contacts dans une Liste de contacts :

```
insert into ContactListMembers (ContactList_ID,Contact_ID) values
('<contactlist_id>','<contact_id>')
```

- Utilisez l'instruction suivante pour supprimer des contacts d'une Liste de contacts :

```
delete from ContactListMembers where ContactList_ID =
'<contactlist_id>' and Contact_ID = '<contact_id>'
```

Restrictions

Voici les restrictions :

- Créer/Mettre à jour :**

- Insérer/Upsert échoue si des colonnes Non nul sont vides ou n'existent pas.
- Insérer/Upsert échoue si les valeurs des colonnes Unique ne sont pas uniques pour un lot donné.
- Afin d'éviter une exception de retour, conservez la valeur de **Nombre de lots à valider** sur 1.


2. Lire :

- a. Pour les entités personnalisées, l'opération `Sélectionner` s'applique uniquement aux jointures avec l'entité `Contacts`.

3. Filtrer :

- a. Les filtres pris en charge sont `=`, `! =`, `>`, `<`, `> =`, `< =`
- b. Les opérateurs de condition `IN` et `NOT IN` ne sont pas pris en charge lors de la fourniture de plus d'une valeur.
- c. `Jointures` n'est pas pris en charge entre les entités.
- d. L'opérateur conditionnel `OR` est pris en charge uniquement pour les entités `Accounts` et `Contacts`.
- e. L'opérateur conditionnel **AND** peut être utilisé uniquement entre deux conditions.
- f. Le filtre `=` ne fonctionne pas toujours sur les champs présentant un type de données `Horodatage`.

Suppression d'une connexion

1. Ouvrez `Management Console`.
2. Accédez à **Ressources > Sources de données**.
3. Cochez la case en regard de la connexion que vous souhaitez supprimer, puis cliquez sur le bouton `Supprimer` .

Exemple de modèle

Introduction

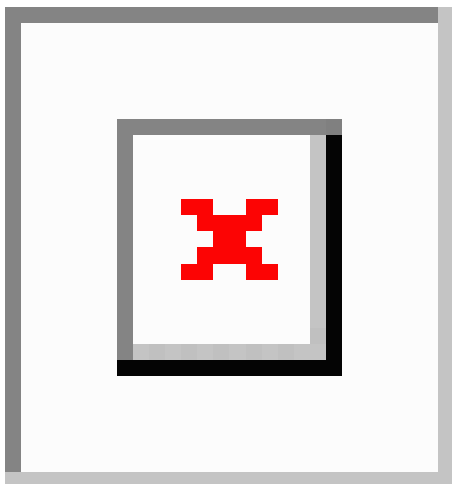
Un exemple de modèle est un flux de données préconfiguré qui permet l'intégration de données entre systèmes avec des connaissances techniques minimales. Vous pouvez utiliser ces flux pour démontrer les fonctionnalités de `Spectrum Technology Platform` avec peu d'efforts. Il n'est pas nécessaire de créer tous les flux de données à partir de zéro. Vous devez créer plusieurs connexions et importer ces flux dans votre système. En quelques minutes, vous pouvez démontrer les fonctionnalités de migration des données d'un système à un autre. Cette section décrit le processus complet de migration des données à l'aide d'exemples de flux.

Remarque : Ces exemples de modèles font partie du fichier du programme d'installation Spectrum et ne fonctionnent pas de manière autonome. Plus aucune assistance n'est disponible pour ces exemples.

Intégration entre SugarCRM OnPremises et Microsoft Dynamics 365 Online

Cette section illustre la migration de Accounts et Contacts de SugarCRM OnPremises vers Microsoft Dynamics 365 Online. S'il existe une association de Contacts et de Accounts dans SugarCRM, la migration fait en sorte de la conserver également dans le système Microsoft Dynamics 365 Online.

Le diagramme suivant illustre le flux de Accounts et Contacts de SugarCRM vers Microsoft Dynamics :



Localisation de l'exemple de modèle

L'exemple de modèle est livré avec le zip du programme d'installation de Spectrum Technology Platform. Suivant votre répertoire d'installation, vous pouvez trouver l'exemple de modèle dans Program Files\Pitney

Bowes\Spectrum\server\modules\metadata\insights\connectors\samples\MS_Dynamics_365_SugarCRM_Accounts_Sync_Dataflow.
Le dossier contient tous les fichiers requis pour configurer le flux de données. Le tableau suivant répertorie les fichiers contenus dans le dossier :

Type	Nom de fichier
Flux de données	<ul style="list-style-type: none"> SugarCRMAccount_Sync_MSDAccount.df SugarCRMContact_Sync_MDSContact.df

Type	Nom de fichier
DBConnection	<ul style="list-style-type: none"> MSDynamics_MS.json SugarCRM_MS.json
ModelStore	<ul style="list-style-type: none"> mi_modelStore_MSDynamics_MS.smims mi_modelStore_SugarCRM_MS.smims

Pré-requis

Vous devez créer deux connexions pour importer et déployer les fichiers requis pour configurer le flux de données. Pour créer les connexions, procédez comme suit :





1. Installez le dernier serveur Spectrum Platform Server.
2. Ouvrez la page d'accueil de Spectrum Server.
3. Cliquez sur **Outils client de la plateforme** , puis sur **Web**
4. Ouvrez la vue **Management Console**.
5. Cliquez sur le menu **Ressources**, puis sur **Sources de données** et créez deux connexions :
 - a. **SugarCRM_TestConnection** de type SugarCRM. Cliquez sur Test et un message de réussite s'affiche : *La connexion SugarCRM_OnPremises à la source de données a bien été établie.*
 - b. **MSDynamics_TestConnection** de type Microsoft Dynamics 365. Cliquez sur Test et un message de réussite s'affiche : *La connexion Microsoft Dynamics 365 Online à la source de données a bien été établie.*

Pour des informations plus détaillées sur la façon de créer des connexions, reportez-vous aux sections Connexion à SugarCRM et Connexion à Microsoft Dynamics 365 Online.

Importation et déploiement des fichiers sur le serveur

Une fois la connexion en place, vous avez maintenant besoin d'un utilitaire à ligne de commande pour importer et déployer les fichiers sur le serveur.

1. Ouvrez la page d'accueil de Spectrum Platform, cliquez sur **Outils client de la plateforme**, puis sur **Ligne de commande** et téléchargez les fichiers suivants :
 - a. Job Executor, qui est un fichier .jar individuel
 - b. Utilitaire Administration, qui est un fichier zippé (spectrum-cli)
2. Décompressez le fichier spectrum-cli.
3. Copiez tous les fichiers du dossier Exemple de modèle dans le dossier spectrum-cli. Cela vous évite d'avoir à saisir le chemin d'accès complet à chaque fichier avec chaque commande.
4. Exécutez l'utilitaire cli.cmd qui se trouve dans le dossier spectrum-cli-12.1. L'interface de ligne de commande de Spectrum Platform est lancée.

5. Connectez-vous au serveur Spectrum Server à l'aide des commandes suivantes : `connect <nom de serveur>:<port> --u <nom d'utilisateur> --p <mot de passe>`. La connexion réussie est confirmée par le message : `Connecté au serveur<nom de serveur> :< port>`.
6. Saisissez et exécutez les commandes suivantes dans l'ordre indiqué pour importer et déployer exemples de flux de données :
 - a. Importez le Model Store à l'aide de la commande : `spectrum> modelstore bulkImport --importDependency true`. Cette commande importe le Model Store avec toutes les dépendances .
 - b. Déployez le Model Store à l'aide de la commande : `modelstore deploy --n <nom>` .
 - c. Importez dbconnection à l'aide de la commande : `dbconnection import --f <nom.json>` .
 - d. Importez les flux de données à l'aide de la commande : `dataflow import --f <nom.df>` .

Remarque :

Supprimez les espaces supplémentaires autour des commandes des noms de fichiers.

Remplacez l'espace réservé <nom> par le nom exact du fichier.

Affichage des fichiers importés

Une fois l'importation terminée, vous pouvez utiliser ces flux de données pour démontrer les fonctionnalités de Spectrum. Les fichiers sont stockés aux emplacements suivants :

1. **Model Store** : dans les onglets Physical Model et Model Store de Metadata Insights
2. **Flux de données** : dans l'Explorateur Serveur de l'application Enterprise Designer
3. **JSON** : dans le Model Store de Management Console

Exécution de la migration des données

Une fois l'importation de données et le déploiement terminés, vous pouvez utiliser les flux de données à l'aide des connexions créées pour migrer les données. Procédez comme suit :

1. Lancez l'application Enterprise Designer. Vous pouvez télécharger le programme d'installation exécutable de cette application depuis la section **Bureau de Outils client de la plateforme** sur la page d'accueil de Spectrum Platform.
2. Connectez-vous à l'aide de vos informations d'identification Spectrum.
3. Cliquez sur Afficher dans le menu, puis sur **Explorateur Serveur**
4. Double-cliquez d'abord sur le job de flux de données **SugarCRMAccount_Sync_MSDAccount**.
5. Double-cliquez sur le stage **Read from DB_SugarCRM**.
 - a. Modifiez la valeur du champ Date saisie, si nécessaire, puis cliquez sur OK.

```
Select "SugarCRM_PM"."Accounts"."email1",
       "SugarCRM_PM"."Accounts"."name",
       "SugarCRM_PM"."Accounts"."phone_office",
```

```
"SugarCRM_PM"."Accounts"."date_entered" From
"SugarCRM_PM"."Accounts" Where
"SugarCRM_PM"."Accounts"."date_entered" Like '2017-08-28%'
```

- b. Cliquez sur **OK** pour continuer.
6. Exécutez le flux en cliquant sur le bouton **Exécuter** de la barre d'outils d'Enterprise Designer.
7. Accédez à la page d'accueil MS Dynamics et cliquez sur **Accounts** sous l'onglet **Customers** sur la page **Sales**. Notez que **Accounts** a été migré.
8. Double-cliquez sur le job **SugarCRMContact_Sync_MDSContacts**.
9. Double-cliquez sur le stage **Read from_MS Dynamics_DB**.
 - a. Modifiez la valeur du champ Créé le, si nécessaire, puis cliquez sur **OK**.

```
Select "MSDynamics_PM"."account"."name",
"MSDynamics_PM"."account"."telephone1",
"MSDynamics_PM"."account"."emailaddress1",
"MSDynamics_PM"."account"."createdon",
"MSDynamics_PM"."account"."accountid" From
"MSDynamics_PM"."account"
Where "MSDynamics_PM"."account"."createdon" Like '2017-09-11%'
```

- b. Cliquez sur **OK** pour continuer.
 10. Exécutez le flux en cliquant sur le bouton **Exécuter** de la barre d'outils d'Enterprise Designer.
 11. Accédez à la page d'accueil MS Dynamics et cliquez sur **Contacts** sous l'onglet **Customers** de la page **Sales**. Notez que les **Contacts** s'affichent sur la page.
- L'affichage des comptes et des contacts sur la page MS Dynamics termine le processus de migration.

2 - Modeling

La production d'informations basées sur des données nécessite des équipes techniques et des équipes commerciales qu'elles aient une compréhension commune des actifs de données de l'entreprise et qu'elles sachent comment ces actifs seront utilisés pour faciliter la prise de décisions commerciales. Les équipes techniques comprennent la conception des bases de données, tandis que les équipes commerciales comprennent les objets d'entreprise (tels qu'un client, un fournisseur ou un magasin) présentant un intérêt. Metadata Insights permet de combler l'écart en fournissant des outils de modeling physique et logique visuellement riches et indépendants l'un de l'autre. Ces outils vous permettent de créer à la fois une vue technique des actifs de données et une vue commerciale des objets d'intérêt, et de relier les deux par mappage.

Un *Physical Model* organise les actifs de données de votre entreprise de manière significative. Un Physical Model permet d'extraire les données à partir de tables, de colonnes et de vues individuelles pour créer une ressource unique que vous pouvez ensuite utiliser pour fournir des données à des Logical Models ou pour effectuer un profilage.

Un *Logical Model* définit les objets qui intéressent votre entreprise et les attributs de ces objets, ainsi que la manière dont ces objets sont reliés entre eux. Par exemple, un Logical Model pour un client peut contenir des attributs de nom et de date de naissance. Il peut également avoir une relation avec un objet d'adresse de domicile, qui contient des attributs de lignes d'adresse, de ville et de code postal. Une fois que vous avez défini les attributs des objets qui intéressent votre entreprise, vous pouvez mapper des sources de données physiques aux attributs du Logical Model, améliorant ainsi l'identification des actifs de données spécifiques qui seront utilisés pour renseigner cet attribut.

Une fois les modèles créés, vous pouvez créer un *Model Store*, qui est une agrégation virtuelle des données référencées par les Physical Models et les Logical Models. Les Model Stores permettent la fédération de données de différentes sources telles que des bases de données, des serveurs de fichiers et des services Cloud. Lorsque vous créez un Model Store, les données des colonnes et des tables qui sont utilisées par le modèle sont extraites et agrégées dans Spectrum™ Technology Platform.

In this section

Logical Model	71
Modèle physique	74
Mapping	82
Model Store	98

Logical Model

Un Logical Model est une représentation d'entités spécifiques à l'entreprise. Il définit les objets qui intéressent votre entreprise et les attributs de ces objets, ainsi que la manière dont ces objets sont reliés entre eux.


Création d'un modèle logique


Un Logical Model est créé par l'ajout d'entités commerciales au modèle et par la définition de relations entre les entités. Vous pouvez également définir les propriétés des entités ainsi que les relations. Lorsque vous avez terminé, vous pouvez mapper le Logical Model vers un ou plusieurs Physical Models.

1. Dans un navigateur Web, accédez à :

`http://server:port/metadata-insights`

Où : *server* est l'adresse IP ou le nom de serveur de votre serveur Spectrum™ Technology Platform et *port* le port HTTP. Par défaut, le port HTTP est 8080.

2. Connectez-vous à l'aide de vos informations d'identification.
3. Cliquez sur **Modeling**.
L'onglet **Logical Model** s'affiche par défaut.
4. Cliquez sur le bouton **Ajouter un modèle logique** .

Remarque : Vous pouvez également choisir de créer un modèle en copiant un modèle existant et en le renommant. Pour ce faire, dans l'onglet **Modèles**, sélectionnez le modèle et cliquez sur l'icône **Copier un modèle** . Renommez le modèle sur la page **Créer Logical Model**, si nécessaire, puis cliquez sur le bouton **Enregistrer**. Pour plus d'informations sur la modification d'un Physical Model, reportez-vous à la section [Modification d'un modèle logique](#) à la page 72.

5. Saisissez un nom pour le modèle dans le champ **Nom**.
6. Dans **Palette d'entités**, faites glisser l'entité que vous souhaitez ajouter au modèle et déposez-la sur le canevas.


Conseil : Utilisez la liste déroulante de la palette pour trier les entités en fonction de leurs catégories. Pour ce faire, vous pouvez également utiliser la zone de filtre à saisie anticipée sous la liste déroulante.

L'entité est ajoutée au canevas et ses propriétés sont affichées dans le volet **Propriétés** sur le côté droit de la page. Pour définir les propriétés d'une entité, reportez-vous à la section [Définition des propriétés d'une entité et d'une relation](#) à la page 72.

7. Pour créer une relation entre deux entités, placez le curseur sur l'entité, glissez la souris et déposez-la sur l'entité à laquelle vous souhaitez la lier.

La relation entre les entités est indiquée par un lien de relation. Lorsque vous placez le curseur de la souris sur l'entité principale, l'entité associée est mise en surbrillance.

Remarque : Comme pour une entité, vous pouvez ajouter des propriétés à la relation à l'aide du volet **Propriétés**. Pour plus d'informations, reportez-vous à la section [Définition des propriétés d'une entité et d'une relation](#) à la page 72.

8. Pour supprimer une entité ou une relation, sélectionnez-la, puis cliquez sur l'icône Supprimer .


Si vous supprimez une entité relationnelle, les liens relationnels associés sont également supprimés.

9. Cliquez sur **Enregistrer**.

Définition des propriétés d'une entité et d'une relation

La boîte de dialogue **Propriétés** affiche le nom par défaut de toutes les relations et entités commerciales. À l'aide de cette boîte de dialogue, vous pouvez modifier le nom et ajouter des attributs aux entités et aux relations d'un Logical Model.

1. Faites un clic droit sur les entités ou les relations dont vous souhaitez définir les propriétés, puis cliquez sur l'option **Propriétés**.
2. Utilisez la section **Ajouter des propriétés** pour ajouter des attributs et leurs types de données à l'entité ou à la relation sélectionnée. Saisissez le nom d'attribut dans le champ **Nom** et sélectionnez son type de données dans la liste déroulante **Type**.
3. Cliquez sur le bouton **Ajouter**.
L'attribut et le type de données définis sont affichés sous les champs d'entrée.

Remarque : Vous pouvez supprimer les détails de l'attribut en cliquant sur l'icône Supprimer  adjacent.

Modification d'un modèle logique

Vous pouvez modifier les détails d'un Logical Model à l'aide de la page **Modifier le Logical Model - <nom du modèle>**.


1. Dans un navigateur Web, accédez à :

`http://server:port/metadata-insights`

Où : *server* est l'adresse IP ou le nom de serveur de votre serveur Spectrum™ Technology Platform et *port* le port HTTP. Par défaut, le port HTTP est 8080.

2. Connectez-vous à l'aide de vos informations d'identification.
3. Cliquez sur **Modeling**.
L'onglet **Logical Model** s'affiche par défaut.
4. Cochez la case en regard du modèle que vous souhaitez modifier.

Remarque : Vous pouvez rechercher un modèle spécifique à l'aide du champ **Filtre** tout en haut de la page. Il recherche les modèles par leur nom.

5. Cliquez sur l'icône Modifier un modèle .
6. Modifiez les détails, selon vos besoins.
7. Cliquez sur **Enregistrer**.

Le modèle sélectionné est mis à jour et apparaît dans la liste de l'onglet **Modèles** de la page **Modeling**.

Suppression d'un modèle logique

Si vous supprimez un modèle logique qui dispose d'un modèle physique dépendant, le modèle physique dépendant devient inutilisable.


1. Dans un navigateur Web, accédez à :

`http://server:port/metadata-insights`

Où : *server* est l'adresse IP ou le nom de serveur de votre serveur Spectrum™ Technology Platform et *port* le port HTTP. Par défaut, le port HTTP est 8080.

2. Connectez-vous à l'aide de vos informations d'identification.
3. Cliquez sur **Modeling**.
L'onglet **Logical Model** s'affiche par défaut.
4. Cochez la case en regard du modèle que vous souhaitez modifier.

Remarque : Vous pouvez rechercher un modèle spécifique à l'aide du champ **Filtre** tout en haut de la page. Il recherche les modèles par leur nom.


5. Cliquez sur le bouton Supprimer un modèle .
6. Cliquez sur **Oui** pour confirmer.

Modèle physique

Un *Physical Model* organise les actifs de données de votre entreprise de manière significative. Un Physical Model permet d'extraire les données à partir de tables, de colonnes et de vues individuelles pour créer une ressource unique que vous pouvez ensuite utiliser pour fournir des données à des Logical Models ou pour effectuer un profilage.

Ajout d'un modèle de données physique


Les Physical Models sont créés en sélectionnant les tables et les colonnes à inclure à partir de la source de données. Vous pouvez créer plusieurs Physical Models qui utilisent les données de la même source de données, mais chaque Physical Model ne peut disposer que d'une seule source de données. Par exemple, un modèle destiné à un administrateur système inclut les tables système de la source de données, tandis que le modèle créé pour les utilisateurs finaux comporte les tables de données requises de la même source de données.


1. Si vous ne l'avez pas déjà fait, utilisez **Management Console** pour connecter Spectrum™ Technology Platform aux sources de données que vous souhaitez modéliser. Pour plus d'informations, reportez-vous à la section [Data Source Connections](#) à la page 9.
Pour accéder à la page **Ajouter une source de données** de **Management Console**, cliquez sur le lien **Créer une connexion**  sur la page.

2. Dans un navigateur Web, accédez à :

```
http://server:port/metadata-insights
```

Où : *server* est l'adresse IP ou le nom de serveur de votre serveur Spectrum™ Technology Platform et *port* le port HTTP. Par défaut, le port HTTP est 8080.

3. Connectez-vous à l'aide de vos informations d'identification.
4. Cliquez sur **Modélisation** > **Physical Model**.
5. Cliquez sur le bouton **Ajouter un modèle physique** .

Remarque : Vous pouvez également choisir de créer un modèle en copiant un modèle existant et en le renommant. Pour ce faire, dans l'onglet **Modèles**, sélectionnez le modèle et cliquez sur l'icône **Copier un modèle** . Renommez le modèle sur la page **Copier un Physical Model**, si nécessaire, puis cliquez sur le bouton **Enregistrer**. Pour plus d'informations sur la modification d'un Physical Model, reportez-vous à la section [Modification d'un modèle physique](#) à la page 75.

6. Dans le champ **Nom**, saisissez un nom pour le modèle.

7. Dans le champ **Type de connexion source** de champ, sélectionnez le type de source de données que vous souhaitez modéliser, puis sélectionnez la source de données spécifique dans le champ **Nom de connexion source**.

Le champ **Nom de connexion source** affiche les sources de données définies dans Management Console pour le type de connexion que vous sélectionnez. Par exemple, si vous sélectionnez Apache Cassandra dans le champ **Type de connexion source**, le champ **Nom de connexion source** répertorie toutes les sources de données Apache Cassandra définies dans Management Console. Si vous souhaitez créer une nouvelle connexion pour ce Physical Model, cliquez sur le lien Créer une connexion en haut à droite de la page. La page **Ajouter une source de données** s'affiche, dans laquelle vous pouvez créer la connexion, l'enregistrer, puis l'utiliser sur cette page.

Remarque : Un Physical Model ne peut disposer que d'une seule source de données.

8. Cliquez sur **Récupérer les métadonnées**.
9. Sélectionnez les données que vous souhaitez inclure.

Pour sélectionner une table ou une vue, cochez la case de la colonne **Inclure** de la table ou de la vue que vous souhaitez inclure. Pour sélectionner toutes les tables répertoriées sur la page, cochez la case **Inclure** dans la ligne d'en-tête. Lorsque vous sélectionnez une table, toutes ses colonnes sont incluses dans le modèle.

Pour sélectionner des colonnes spécifiques à inclure dans le modèle, passez la souris sur la table contenant les colonnes souhaitées, puis cliquez sur le bouton **Sélectionner des colonnes** qui s'affiche.

Vous pouvez rechercher des tables en fonction de leur nom ou de leur type à l'aide du champ **Filtre** au-dessus de la liste des tables. Par exemple, si vous souhaitez afficher uniquement les tables dont les noms contiennent le terme « nouveau », saisissez `nouveau` dans le champ Filtre. De même, vous pouvez filtrer les tables en fonction de leur type. Cependant, vous ne pouvez utiliser qu'un seul des critères de filtrage, soit Nom, soit Type.

10. Cliquez sur **Enregistrer**.

Le modèle est créé et apparaît dans la liste de l'onglet **Modèles** de la page **Modeling**.

Modification d'un modèle physique

1. Dans un navigateur Web, accédez à :

`http://server:port/metadata-insights`

Où : *server* est l'adresse IP ou le nom de serveur de votre serveur Spectrum™ Technology Platform et *port* le port HTTP. Par défaut, le port HTTP est 8080.


2. Connectez-vous à l'aide de vos informations d'identification.

3. Cliquez sur **Modélisation > Physical Model**.
4. Cochez la case en regard du modèle que vous souhaitez modifier.

Remarque : Vous pouvez rechercher un modèle spécifique à l'aide du champ **Filtre** tout en haut de la page. Il recherche les modèles par leur nom.

5. Cliquez sur le bouton Modifier un modèle .

La page **Modifier un Physical Model** s'affiche pour le modèle sélectionné avec la case à cocher **Inclure** affichant les tables et les colonnes déjà sélectionnées.

Remarque : Les Physical Models sont affichés dans deux vues : Liste et sous forme de tableau. La vue Liste affiche la table sélectionnée sous forme de liste sous les champs de spécification de connexion. La vue sous forme de tableau affiche uniquement les tables sélectionnées et leurs détails. Pour utiliser la vue sous forme de tableau, cliquez sur l'icône de la vue sous forme de tableau  en haut de la page. Pour plus d'informations, reportez-vous à la section **Vue sous forme de tableau d'un Physical Model** à la page 76.

6. Modifiez la connexion du modèle, si nécessaire, à l'aide de la liste déroulante **Nom de connexion source**.
7. Supprimez la table ajoutée, ou ajoutez une table, si nécessaire, en cliquant sur les lignes correspondantes.
8. Cliquez sur **Enregistrer**.

Le modèle sélectionné est mis à jour et apparaît dans l'onglet **Modèles** de la page **Modeling**.

Vue sous forme de tableau d'un Physical Model

La vue sous forme de tableau d'un Physical Model affiche les détails de toutes les tables de données sélectionnées. Chaque table affiche les attributs et son type de données. D'autres propriétés des tables sont affichées dans le volet **Propriétés**.

1. Dans un navigateur Web, accédez à :


`http://server:port/metadata-insights`

Où : *server* est l'adresse IP ou le nom de serveur de votre serveur Spectrum™ Technology Platform et *port* le port HTTP. Par défaut, le port HTTP est 8080.

2. Cliquez sur **Modélisation > Physical Model**.
3. Cochez la case en regard du modèle que vous souhaitez afficher.

Remarque : Vous pouvez rechercher un modèle spécifique à l'aide du champ **Filtre** tout en haut de la page. Il recherche les modèles par leur nom.

4. Cliquez sur l'icône Afficher un modèle .

Remarque : Vous pouvez également accéder à la vue sous forme de tableau en cliquant sur l'icône de la vue sous forme de tableau  de la page **Modifier un Physical Model**.

5. Pour afficher les propriétés d'une table, sélectionnez-la, puis cliquez sur l'onglet réductible **Propriétés** sur le côté droit de la page.

Remarque : Vous ne pouvez pas modifier ces propriétés.

Le volet **Propriétés** affiche ces détails :

Remarque : Les détails affichés ici sont basés sur le type de la source de données sélectionnée. Pour plus d'informations, reportez-vous à la section [Détails spécifiques à la connexion](#) à la page 77.

- **Colonnes** : détails de la colonne, tels que le nom de la colonne, le type de données qu'elle contient, sa taille et de si elle est annulable et actualisable.
- **Clés primaires** : nom de la clé primaire (si définie) et colonne et séquence de clé associées
- **Clés étrangères** : nom de la clé étrangère (si définie) et colonne et nom de table associés

Détails spécifiques à la connexion

Les propriétés de la table, telles que les colonnes, les clés primaires et les clés étrangères, affichées dans l'onglet réductible **Propriétés** de la page **Modifier un Physical Model**, sont basées sur le type de connexion source sélectionné dans le Physical Model. Les détails des différents types de connexion sont fournis ci-dessous.

Remarque : Pour connaître les étapes permettant d'afficher les **Propriétés** des tables, reportez-vous à la section [Vue sous forme de tableau d'un Physical Model](#) à la page 76.

Connexion JDBC	Colonnes	Affiche le nom, le type de données et la taille de chaque colonne. La section indique également si une colonne peut être annulée et/ou s'il s'agit d'une clé primaire.
	Clés primaires	Affiche la clé primaire, le nom de clé primaire et la séquence de clé.
	Clés étrangères	Affiche le nom de colonne de clés étrangères, le nom de clé étrangère, le nom de catalogue de clés primaires, le nom de table de clés primaires, le nom de colonne de clés primaires et le nom de clé primaire.
	Index	Affiche le nom d'index et le nom de colonne et indique si les index sont uniques ou non. Si un index n'est pas unique, la case NU est cochée ; sinon elle est décochée.
Connexion Amazon DynamoDB	Colonnes	Affiche le nom et le type de données de chaque colonne. La section indique également si une colonne peut être mise à jour, faire l'objet d'une recherche et/ou s'il s'agit d'une clé primaire.

	Clés primaires	Affiche la clé primaire et le nom de clé primaire.
Connexion Amazon SimpleDB	Colonnes	Affiche le nom de colonne et le type de données de la table que vous sélectionnez.
Connexion Apache Cassandra	Colonnes	Affiche le nom de colonne et le type de données. La section indique également si une colonne peut être mise à jour, faire l'objet d'une recherche et/ou s'il s'agit d'une clé primaire.
	Clés primaires	Affiche la clé primaire et le nom de clé primaire.
Connexion Flat File	Colonnes	Affiche le nom de colonne et le type de données de la table que vous sélectionnez.
Connexion Marketo	Colonnes	Affiche le nom de colonne et le type de données. La section indique également si une colonne est une clé primaire, accessible en lecture ou capable d'être mise à jour.
	Clé principale	Affiche la clé primaire et le nom de clé primaire.
	Clés étrangères	Affiche le nom de colonne de clés étrangères, le nom de clé étrangère et le nom de table de référence.
	Droits	Indique les fonctions pouvant être exécutées sur la table Physical Model sélectionnée. Les fonctions Lire, Insérer, Mettre à jour, Upsert et Supprimer peuvent être réalisées sur des tables Marketo.
	Types de filtre	Indique la colonne sur laquelle la recherche est applicable. Par exemple, <i>id</i> comme Type de filtre dans Prospect indique que Prospect peut faire l'objet d'une recherche à l'aide de <i>id</i> .

Remarque : Marketo ne prend pas en charge plusieurs types de filtre dans la requête `Select`.

Par exemple :

Requête valide :

```
Select * from Lead where id = 1234
```

```
Select * from Lead where email = 'abc@txyz.com'
```

Requête non valide :

```
Select * from Lead where id = 1234 and email = 'abc@txyz.com'
```

Connexion MS Dynamics 365 Online	Colonnes	Affiche le nom de colonne et le type de données. La section indique également si une colonne peut être annulée et/ou s'il s'agit d'une clé primaire.
	Clés primaires	Affiche la clé primaire et le nom de clé primaire.
	Clés étrangères	Affiche le nom de colonne de clés étrangères, le nom de clé étrangère et le nom de table de référence.
	Droits	Indique les fonctions pouvant être exécutées sur la table Physical Model sélectionnée. Actuellement, les fonctions Créer, Lire, Écrire et Supprimer peuvent être réalisées sur une table MS Dynamics 365 Online.
Connexion NetSuite	Colonnes	Affiche le nom, le type de données et la taille de chaque colonne. La section indique également si une colonne peut être Null et/ou une clé principale, et quelles fonctions Créer, Lire, et Mettre à jour peuvent être réalisées sur la colonne.
	Clés primaires	Affiche la clé primaire et le nom de clé primaire.
Connexion Salesforce	Colonnes	Affiche le nom, le type de données et la taille de chaque colonne. La section indique également si une colonne peut être annulée et/ou s'il s'agit d'une clé primaire.
	Clés primaires	Affiche la clé primaire et le nom de clé primaire.
	Clés étrangères	Affiche le nom de colonne de clés étrangères, le nom de clé étrangère et le nom de table de référence.
Connexion SAP NetWeaver	Colonnes	Affiche le nom, le type de données et la taille de chaque colonne. La section indique également si une colonne peut être annulée et/ou s'il s'agit d'une clé primaire.
	Clés primaires	Affiche la clé primaire et le nom de clé primaire.
	Clés étrangères	Affiche le nom de colonne de clés étrangères, le nom de clé étrangère et le nom de table de référence.
	Droits	Indique les fonctions pouvant être exécutées sur la table Physical Model sélectionnée. Actuellement, les fonctions Créer, Écrire et Supprimer peuvent être réalisées sur une table SAP NetWeaver.
Connexion Splunk	Colonnes	Affiche le nom et le type de données de chaque colonne. La section indique également si une colonne peut être mise à jour, faire l'objet d'une recherche et/ou s'il s'agit d'une clé primaire.
	Clés primaires	Affiche la clé primaire et le nom de clé primaire.
Connexion SuccessFactors	Colonnes	Affiche le nom, le type de données et la taille de chaque colonne. La section indique également si une colonne peut être Null et/ou une clé principale, et quelles fonctions Filtrer, Insérer, Trier, Mettre à jour, Upsert, et Afficher peuvent être réalisées sur la colonne.

	Clés primaires	Affiche les colonnes de clé primaire et le nom de clé primaire respectif dont chacune fait partie.
	Clés étrangères	Affiche le nom de colonne de clés étrangères, le nom de clé étrangère et le nom de table de référence.
	Droits	Indique lesquelles parmi les fonctions Insérer, Mettre à jour, Upsert et Supprimer peuvent être réalisées sur la table Physical Model sélectionnée.
Connexion SugarCRM	Colonnes	Affiche le nom, le type et la taille de chaque colonne. Indique également si une colonne est annulable et actualisable.
	Clés primaires	Affiche les noms de clé primaire, la colonne et la séquence de clés.
	Clés étrangères	Affiche le nom de colonne de clés étrangères, le nom de clé étrangère et le nom de table de référence.
Connexion Siebel	Colonnes	Affiche le nom, le type et la taille de chaque colonne. La section indique également si une colonne est annulable et actualisable.
	Clés primaires	Affiche les noms de clé primaire, la colonne et la séquence de clés.
	Clés étrangères	Affiche le nom de colonne de clés étrangères, le nom de clé étrangère et le nom de table de référence.

Restrictions de lecture et d'écriture

Restriction de lecture

- Metadata Insights n'interprète pas les noms de champ sensibles à la casse comme des noms différents. Si deux tables ou deux colonnes portent le même nom, mais avec une casse différente, elles sont interprétées comme des noms doublons. Si un Model Store contient de tels noms de champ, son déploiement échoue.

Par exemple, des champs comme `DeDup`, `Dedup` et `DEdup` sont interprétés comme des noms doublons. Le déploiement d'un Model Store contenant de tels champs avec une casse différente échoue.

- Si une Data Source, un Physical Model, un Logical Model ou un Model Store dépend d'une source de données JDBC spécifique, il/elle ne peut être ni importé(e), ni exporté(e) via l'utilitaire Administration.

Restrictions d'écriture

Lors de l'écriture dans une source de base de données à l'aide du stage **Write to DB** d'**Enterprise Designer**, il existe certaines restrictions.

- La création de tables lors de l'écriture vers une source de base de données à l'aide d'une connexion Metadata Insights n'est pas prise en charge. Si vous cliquez sur **Créer une table** tout en utilisant une connexion Metadata Insights, une exception d'exécution s'affiche.
- Metadata Insights ne prend pas en charge les écritures unifiées simultanées vers plusieurs connexions. Vous ne pouvez écrire qu'une seule source à la fois.
- L'abandon et la recréation d'une table tout en écrivant vers une base de données à l'aide d'une connexion Metadata Insights ne sont pas pris en charge.
- Les opérations de suppression et de tronquage ne sont pas prises en charge lors de l'écriture vers une base de données via Metadata Insights. Seules les opérations d'insertion et de mise à jour sont prises en charge.
- Metadata Insights procède à une validation à la fin de chaque insertion de lot. Le champ **Nombre de lots à valider** ne peut être modifiée sur aucune autre valeur que 1.

Dans l'onglet **Exécution** du stage **Write to DB** d'**Enterprise Designer**, le champ **Nombre de lots à valider** ne peut comporter que la valeur 1 pour Metadata Insights.

Remarque : Par défaut, un lot comporte 1 000 enregistrements. Vous pouvez configurer la taille de lot en fonction de votre système, dans l'onglet **Exécution** du stage **Write to DB** d'**Enterprise Designer**.

- Pour les insertions de lots, la fonction **Valider à la fin** n'est pas prise en charge. Cette fonction fait référence à une validation une fois l'ensemble des lots terminés. Dans Metadata Insights, une validation est effectuée à la fin de chaque insertion de lot.

Suppression d'un Physical Model

Suivez cette procédure pour supprimer un Physical Model de Metadata Insights.

Remarque : Si vous supprimez un Physical Model mappé vers un Logical Model, le Logical Model devient inutilisable.

1. Dans un navigateur Web, accédez à :

`http://server:port/metadata-insights`

Où : *server* est l'adresse IP ou le nom de serveur de votre serveur Spectrum™ Technology Platform et *port* le port HTTP. Par défaut, le port HTTP est 8080.

2. Connectez-vous à l'aide de vos informations d'identification.
3. Cliquez sur **Modélisation > Physical Model**.
4. Cochez la case en regard du modèle que vous souhaitez supprimer.

Remarque : Vous pouvez rechercher un modèle spécifique à l'aide du champ **Filtre** tout en haut de la page. Il recherche les modèles par leur nom.

5. Cliquez sur le bouton Supprimer un modèle .

6. Cliquez sur **Oui** pour confirmer.

Mapping

Mappage de modèles

Les Logical Models doivent être mappés vers un ou plusieurs modèles de données physiques pour définir les sources de données à utiliser pour renseigner les entités du Logical Model. Ce mappage peut être effectué à l'aide d'une ou de plusieurs de ces méthodes :

- Directement : faites glisser une table du Physical Model jusqu'au canevas et établissez un lien entre les colonnes des tables logiques et physiques.
- Transformation : transformez les valeurs avant de les fournir aux tables Logical Model. La transformation peut être l'application d'une fonction de conversion ou de fonctions string, numeric et datetime à une colonne Logical Model. Elle peut également impliquer l'application d'une valeur de constante à une colonne d'une table Logical Model.

-

Vous pouvez filtrer les données à fournir aux entités Logical Model :

- Génération de critères : définissez un critère de filtrage des données extraites des colonnes Physical Model. Les données filtrées sont chargées dans le Logical Model.
- Jointure de colonnes et de tables : combinez des données des colonnes de deux ou plusieurs tables Physical Model avant de les fournir aux colonnes Logical Model.



Si vous mappez des colonnes de deux types de données différents, le type de données de la colonne Logical Model est mis à jour sur celui de la colonne Physical Model.

Mappage d'entités logiques vers des données physiques

1. Dans un navigateur Web, accédez à :

`http://server:port/metadata-insights`

Où : *server* est l'adresse IP ou le nom de serveur de votre serveur Spectrum™ Technology Platform et *port* le port HTTP. Par défaut, le port HTTP est 8080.

2. Connectez-vous à l'aide de vos informations d'identification.
3. Cliquez sur **Modeling**.
4. Sélectionnez le Logical Model que vous souhaitez mapper, puis cliquez sur l'icône **Mapper modèle**  dans la barre d'outils ou sur l'icône de mappage  de la colonne **Mappé**.

La couleur de l'icône de la colonne **Mappé** indique s'il est mappé. Une couleur verte indique un mappage partiel ou complet, tandis qu'une couleur rouge indique que le modèle n'a pas encore été mappé.

5. Dans le volet **Physical Model**, sélectionnez le Physical Model que vous souhaitez mapper ou sa table ou sa colonne vers laquelle vous souhaitez mapper les entités de votre Logical Model. Il existe plusieurs façons de créer le mappage :


Remarque : Pour créer un Physical Model, cliquez sur le volet et spécifiez les détails dans la fenêtre contextuelle **Créer un Physical Model**. Ce modèle est désormais disponible pour le mappage.

- Pour effectuer le mappage à l'aide de glisser-déposer :
 1. Dans la liste de modèles, cliquez sur le Physical Model requis. Une liste de toutes les tables contenues dans le modèle s'affiche.
 2. Faites glisser la table de votre choix jusqu'au canevas. La table affiche les colonnes et le type de données de chaque colonne.
 3. Mappez une colonne vers un attribut d'une entité logique en glissant la souris de la colonne vers l'attribut.
- Pour effectuer le mappage à l'aide de la zone de recherche (utile lorsque vous connaissez uniquement le nom de table et non le modèle auquel elle appartient) :



1. Saisissez le terme de recherche dans le champ de recherche, et, dans la liste déroulante adjacente, sélectionnez **Colonnes**, **Tables** ou **Tout** suivant que le terme doit figurer uniquement dans des colonnes ou dans des tables, ou dans des colonnes, des tables ainsi que des modèles.

Les résultats sont affichés sous les champs de recherche, indiquant le nom de la colonne, le type de données, le nom de la table ou de la vue et le nom du Physical Model auquel appartient la table.

2. Si vous avez recherché une table, cliquez dessus, puis sur le bouton **Ajouter au canevas**. La table sélectionnée est ajoutée au canevas et vous pouvez la mapper via la méthode glisser-déposer, telle que décrite ci-dessus.
3. Si vous avez recherché une colonne, cliquez sur le nom de colonne. Les champs **Entité** et **Attribut** apparaissent. Le champ **Entité** répertorie toutes les entités du Logical Model actuellement mappé, tandis que le champ **Attribut** répertorie tous les attributs de l'entité sélectionnée.

Sélectionnez l'entité et l'attribut que vous souhaitez mapper, puis cliquez sur l'icône de mappage .

Remarque : Un attribut peut être mappé vers une seule colonne à la fois.

- Pour convertir une valeur de la colonne Physical Model avant de la mapper vers une entité logique, utilisez une transformation. Pour plus d'informations, reportez-vous à la section [Application d'une transformation](#) à la page 85.
 - Pour filtrer les données d'un Logical Model, cliquez sur l'icône Générateur de critères . Pour plus d'informations, reportez-vous à la section [Critères](#) à la page 94.
 - Pour combiner les données des colonnes de deux tables Physical Model avant de renseigner les colonnes Logical Model, reportez-vous à la section [Jointures](#) à la page 96.
6. Pour supprimer l'un des mappages ou l'une des tables du canevas, sélectionnez-le/la, puis cliquez sur l'icône **Supprimer mappage**  de la barre d'outils.
 7. Cliquez sur **Enregistrer**.

Le mappage est enregistré et l'icône de mappage passe du rouge au vert pour ce Logical Model dans l'onglet **Modèles** de la page **Modeling**.

Transformations

Cette fonctionnalité vous permet de transformer une valeur avant de la fournir dans une colonne d'une entité Logical Model. Les valeurs proviennent des colonnes des tables Physical Model.

Suivant les fonctions que vous devez appliquer aux valeurs avant de renseigner les colonnes Logical Model, vous pouvez effectuer une transformations sur :

- Des colonnes : pour fournir les valeurs d'une colonne d'une table Physical Model.
- Une constante : pour appliquer une valeur de constante à une colonne Logical Model ou à l'un des paramètres d'une transformation de fonction.
- Une fonction : pour appliquer des fonctions de chaîne, de valeur numérique, de conversion et de valeur datetime à une colonne Logical Model ou à l'un des paramètres d'une transformation de fonction.

Vous pouvez créer des transformations imbriquées complexes au sein de la transformation parente.

Transformation de fonction de conversion

Une fonction de conversion vous permet de convertir des valeurs d'un type de données en un autre type de données, d'un format en un autre format, ou d'analyser des valeurs en fonction de formats souhaités.

Une fonction de conversion est appliquée pour convertir les données d'une colonne d'une table source ou d'une constante en type de données ou format souhaité. Ces données converties renseignent la colonne sélectionnée d'une entité logique.

Les différentes fonctions de conversion disponibles sont les suivantes :

cast	Cette fonction vous permet de convertir des valeurs d'un type de données en un autre.
-------------	---

format	Cette fonction vous permet de convertir une valeur d'un type de données en une valeur de type chaîne à l'aide d'un format fourni.
parse	Cette fonction vous permet d'analyser une valeur de type chaîne dans le type de données souhaité à l'aide d'un format fourni.
to_bytes	Cette fonction vous permet de convertir une valeur de type CLOB en octets.
to_chars	Cette fonction vous permet de convertir une valeur de type BLOB en caractères.

Paramètres des fonctions de conversion

Les fonctions de conversion `cast`, `format` et `parse` comportent deux paramètres :

- Valeur source d'un certain type de données et d'un certain format
- Type de données ou format cible dans lequel la valeur source doit être convertie ou à l'aide duquel la valeur source doit être analysée.


Remarque : Cette convention n'est pas applicable aux fonctions de conversion `to_bytes` et `to_chars`.

Application d'une transformation

1. Sur la page **Créer mappage <nom de Logical Model>**, sélectionnez l'entité correspondant à l'attribut dont vous souhaitez appliquer la valeur transformée, puis cliquez sur l'onglet réductible **Transformations** sur le côté droite de la page.

Le volet **Transformations** s'affiche, indiquant le nom de tous les attributs de l'entité sélectionnée. La colonne **Transformation** reste initialement vide. Elle affiche le type de transformation une fois que vous avez enregistré la transformation.

Remarque : Le volet **Transformation** n'affiche pas les détails de la table Physical Model. Si vous cliquez sur une table Physical Model après une entité et si vous ouvrez le volet **Transformation**, il affiche les détails de la dernière entité sélectionnée.

2. Sélectionnez l'attribut auquel vous souhaitez appliquer la valeur transformée, puis cliquez sur l'icône **Transformation**  . La fenêtre contextuelle **Transformation de colonne '<nom de colonne>'** apparaît.
3. Sélectionnez une option dans **Colonne**, **Constante** ou **Fonction** pour définir le type de transformation à appliquer sur les valeurs des tables de Physical Models.

Option	Description
Colonne	Dérivez la valeur d'une colonne de table de Physical Models. Pour plus d'informations, reportez-vous à la section Application d'une transformation de colonne à la page 86.

Option	Description
Constante	Définissez la valeur obtenue comme constante. Pour plus d'informations, reportez-vous à la section Application d'une transformation de constante à la page 87.
Fonction	Définissez la valeur obtenue à l'aide des fonctions String, Numeric, Conversion et Datetime. Pour plus d'informations, reportez-vous à la section Application d'une transformation de fonction à la page 87.


- Définissez la transformation à l'aide de l'une des trois options ci-dessus, ou d'une combinaison de votre choix de n'importe lesquelles des trois.
- Cliquez sur **OK**.
La transformation définie est appliquée à l'attribut et s'affiche avec une flèche sur le canevas. Le type de transformation est également affiché dans le volet **Transformations** de la page **Créer mappage**.

[Application d'une transformation de colonne](#)

Une transformation de colonne vous permet de lier la colonne requise d'une table de Physical Models pour définir un attribut d'une entité logique ou l'un des paramètres d'une transformation de fonction. Vous pouvez effectuer une transformation de colonne à l'aide du volet **Transformations** comme décrit ci-dessous, ou en liant une colonne Physical Model à un attribut d'une entité logique directement sur le canevas.

- Sur la page **Créer mappage <nom de Logical Model>**, sélectionnez l'entité correspondant à l'attribut dont vous souhaitez appliquer la valeur transformée, puis cliquez sur l'onglet réductible **Transformations** sur le côté droite de la page.
Le volet **Transformations** s'affiche, indiquant le nom de tous les attributs de l'entité sélectionnée. La colonne **Transformation** reste initialement vide. Elle affiche le type de transformation une fois que vous avez enregistré la transformation.

Remarque : Le volet **Transformation** n'affiche pas les détails de la table Physical Model. Si vous cliquez sur une table Physical Model après une entité et si vous ouvrez le volet **Transformation**, il affiche les détails de la dernière entité sélectionnée.

- Sélectionnez l'attribut auquel vous souhaitez appliquer la valeur transformée, puis cliquez sur l'icône **Transformation**  .
La fenêtre contextuelle **Transformation de colonne '<nom de colonne>'** apparaît.
- Sélectionnez l'option **Colonne** et, dans la liste déroulante **Nom du Physical Model**, sélectionnez le Physical Model requis.
- Dans la liste déroulante **Nom de la table**, sélectionnez la table requise.
- Dans la liste déroulante **Nom de colonne**, sélectionnez la colonne requise que vous souhaitez mapper vers l'attribut.
- Cliquez sur **OK**.

La colonne sélectionnée est mappée vers l'attribut, et cela est indiqué par une flèche de connexion entre les deux sur le canevas. Le volet **Transformation** affiche cela sous forme de transformation **Simple**.


Remarque : Le type de données de l'attribut est hérité de la colonne de Physical Models mappée.


Application d'une transformation de constante

Une transformation de constante vous permet d'appliquer une valeur de constante pour définir un attribut d'une entité logique ou l'un des paramètres d'une transformation de fonction.

1. Sur la page **Créer mappage <nom de Logical Model>**, sélectionnez l'entité correspondant à l'attribut dont vous souhaitez appliquer la valeur transformée, puis cliquez sur l'onglet réductible **Transformations** sur le côté droite de la page.
Le volet **Transformations** s'affiche, indiquant le nom de tous les attributs de l'entité sélectionnée. La colonne **Transformation** reste initialement vide. Elle affiche le type de transformation une fois que vous avez enregistré la transformation.

Remarque : Le volet **Transformation** n'affiche pas les détails de la table Physical Model. Si vous cliquez sur une table Physical Model après une entité et si vous ouvrez le volet **Transformation**, il affiche les détails de la dernière entité sélectionnée.

2. Sélectionnez l'attribut auquel vous souhaitez appliquer la valeur transformée, puis cliquez sur l'icône **Transformation** .
La fenêtre contextuelle **Transformation de colonne '<nom de colonne>'** apparaît.
3. Sélectionnez l'option **Constante** et, dans la liste déroulante **Type**, sélectionnez le type de données de la valeur de constante que vous souhaitez définir.
4. Dans le champ **Valeur**, saisissez une valeur de constante pour le type de données sélectionné.
5. Cliquez sur **OK**.

La transformation de constante appliquée est reflétée sur le canevas à l'aide d'un nœud de transformation  et elle est indiquée comme une transformation de **Constante** dans le volet **Transformation**.

Remarque : La fenêtre contextuelle **Transformation de colonne '<nom de colonne>'** s'affiche lorsque vous double-cliquez sur ce nœud.


Application d'une transformation de fonction

Une transformation de fonction vous permet d'appliquer des fonctions de type chaîne, numérique, conversion et datetime pour définir un attribut d'une entité logique, ou l'un des paramètres d'une transformation de fonction parente.

1. Sur la page **Créer mappage <nom de Logical Model>**, sélectionnez l'entité correspondant à l'attribut dont vous souhaitez appliquer la valeur transformée, puis cliquez sur l'onglet réductible **Transformations** sur le côté droite de la page.


Le volet **Transformations** s'affiche, indiquant le nom de tous les attributs de l'entité sélectionnée. La colonne **Transformation** reste initialement vide. Elle affiche le type de transformation une fois que vous avez enregistré la transformation.

Remarque : Le volet **Transformation** n'affiche pas les détails de la table Physical Model. Si vous cliquez sur une table Physical Model après une entité et si vous ouvrez le volet **Transformation**, il affiche les détails de la dernière entité sélectionnée.

- Sélectionnez l'attribut auquel vous souhaitez appliquer la valeur transformée, puis cliquez sur l'icône **Transformation**  .
La fenêtre contextuelle **Transformation de colonne '<nom de colonne>'** apparaît.
- Sélectionnez l'option **Fonction** et, dans la liste déroulante **Catégorie**, sélectionnez la catégorie de la fonction que vous souhaitez appliquer. Vous pouvez choisir entre **Chaîne**, **Numérique**, **Conversion** et **Datetime**.
- Dans la liste déroulante **Nom**, sélectionnez la fonction que vous souhaitez appliquer.

Remarque : Les options de cette liste déroulante dépendent de la fonction **Catégorie** que vous sélectionnez.

La fonction est définie et s'affiche dans la fenêtre sous les champs d'entrée.

Remarque : Vous pouvez afficher la description de la fonction en passant la souris sur l'icône  sous la liste déroulante **Nom**.

- Définissez les paramètres de la fonction sélectionnée en la développant et en utilisant l'une de ces options : **Colonne**, **Constante** ou **Fonction**.

Remarque : Assurez-vous que le type de données obtenu correspond au type de données prévu du paramètre de la fonction.

- Cliquez sur **OK**.
La transformation de fonction appliquée est reflétée sur le canevas l'aide d'un nœud de transformation.
- Cliquez sur **Enregistrer**.

Le mappage est enregistré et l'icône de mappage passe du rouge au vert pour ce Logical Model dans l'onglet **Modèles** de la page **Modeling**.

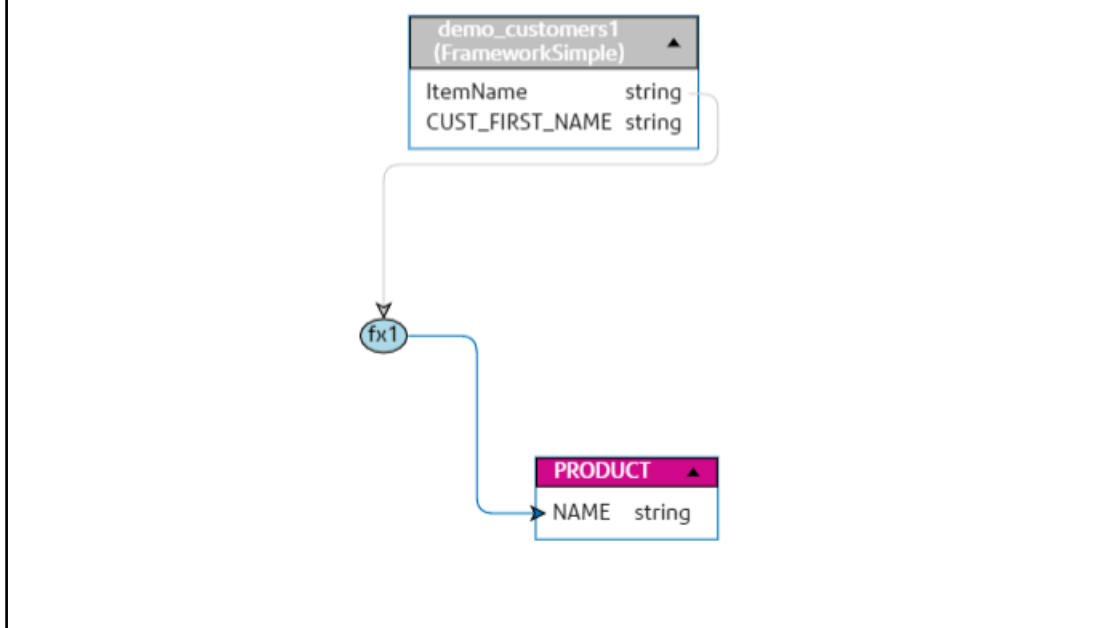
Exemple - Application d'une transformation de fonction à un attribut d'une entité Logical Model

Par exemple, si vous souhaitez supprimer les espaces à droite de chaque valeur de la colonne `ItemName` de la table Physical Model `demo_Customers1` et fournir la valeur supprimée dans la colonne liée `Name` de la table logique `Product` :

- Sélectionnez la table `Product` sur le canevas.

2. Dans le volet **Transformations**, cliquez sur Δ en regard de la colonne `Name`, et, dans la fenêtre contextuelle **Transformation de colonne 'NOM'**, sélectionnez l'option **Fonction**.
3. Dans la liste déroulante **Catégorie**, sélectionnez `String`.
4. Dans la liste déroulante **Nom**, sélectionnez la fonction `rtrim(string string)`.
5. La fonction **Chaîne** est ajoutée à la fenêtre contextuelle sous les champs d'entrée. Cliquez sur cette fonction et définissez le paramètre auquel cette fonction doit être appliquée à l'aide de l'option **Colonne**, de telle sorte que le type de données obtenu du paramètre soit une chaîne.
6. Cliquez sur **OK**.

La transformation de fonction appliquée est reflétée sur le canevas comme suit :



Application d'une transformation de conversion

Remarque : Les étapes ci-dessous ne s'appliquent pas aux fonctions de conversion `to_bytes` ni `to_chars`.

1. Sur la page **Créer mappage <nom de Logical Model>**, sélectionnez l'entité correspondant à l'attribut dont vous souhaitez appliquer la valeur transformée, puis cliquez sur l'onglet réductible **Transformations** sur le côté droite de la page. Le volet **Transformations** s'affiche, indiquant le nom de tous les attributs de l'entité sélectionnée. La colonne **Transformation** reste initialement vide. Elle affiche le type de transformation une fois que vous avez enregistré la transformation.

Remarque : Le volet **Transformation** n'affiche pas les détails de la table Physical Model. Si vous cliquez sur une table Physical Model après une entité et si vous ouvrez le volet **Transformation**, il affiche les détails de la dernière entité sélectionnée.

2. Sélectionnez l'attribut auquel vous souhaitez appliquer la valeur transformée, puis cliquez sur l'icône **Transformation** Δ .
La fenêtre contextuelle **Transformation de colonne** '<nom de colonne>' apparaît.
3. Sélectionnez l'option **Fonction**, et, dans la liste déroulante **Catégorie**, sélectionnez l'option **Conversion**.
4. Dans le champ **Nom**, sélectionnez la fonction de conversion requise.
La fonction est définie comme une conversion et la valeur et la cible s'affichent sous les champs d'entrée.
 - **Valeur** : valeur source à convertir.
 - **Cible** : type de données ou format cible dans lequel la valeur source doit être convertie ou à l'aide duquel la valeur source doit être analysée.
5. Sélectionnez la **Valeur** à l'aide de l'une des options suivantes : **Colonne**, **Constante** ou **Fonction**.

Remarque : Assurez-vous que le type de données obtenu correspond au type de données prévu pour le premier paramètre de la fonction cast (sélectionnée dans le champ **Nom** ci-dessus).
6. Sélectionnez la **Cible** à l'aide de l'une des options suivantes : **Colonne**, **Constante** ou **Fonction**.
Pour plus d'informations, reportez-vous à la section [Types de données pris en charge pour les transformations de conversion](#) à la page 91.
7. Cliquez sur **OK**.
La transformation de fonction de conversion appliquée est reflétée sur le canevas à l'aide d'un lien et du nœud de fonction.
8. Cliquez sur **Enregistrer**.

Le mappage est enregistré et l'icône de mappage passe du rouge au vert pour ce Logical Model dans l'onglet **Modèles** de la page **Modeling**.

Exemple - Application d'une transformation de fonction de conversion cast

Par exemple : pour convertir une valeur entière de la colonne `BALANCE` de la table `CUSTOMER` du Physical Model `ClientSheet` en une valeur à virgule flottante pour renseigner la colonne `CUSTBALANCE` d'une entité logique `CUSTPROFILE` :

1. Sélectionnez l'entité `CUSTPROFILE` sur le canevas.
2. Cliquez sur le volet **Transformations** et, dans le volet, cliquez sur l'icône **Transformation** Δ en regard de la colonne `CUSTBALANCE`.
3. Dans la fenêtre contextuelle, sélectionnez l'option **Fonction**.
4. Dans le champ **Catégorie**, sélectionnez `Conversion`.
5. Dans le champ **Nom** sélectionnez la fonction `cast(Integer value, string target)`. Vous avez la possibilité de définir la valeur et la cible sous les champs d'entrée.
6. Cliquez sur **Valeur** et sélectionnez l'option **Colonne**.

- Dans la liste déroulante **Nom du Physical Model**, sélectionnez le Physical Model `ClientSheet`.
- Dans la liste déroulante **Nom de table**, sélectionnez une table `CUSTOMER`.
- Dans la liste déroulante **Nom de colonne**, sélectionnez la colonne `BALANCE`.

La valeur Entier de cette colonne devient le premier paramètre de la fonction `cast`.

7. Cliquez sur **Cible** et effectuez ces sélections.

Remarque : Dans cet exemple, le type de données cible prévu est une chaîne.

- Sélectionnez l'option **Constante**.
- Dans la liste déroulante **Type**, sélectionnez le type de données `String`.
- Dans le champ **Valeur**, saisissez `Flottant`. Il s'agit du type de données cible.

Le type de données cible `Flottant` devient le deuxième paramètre de la fonction `cast`.

8. Cliquez sur **OK**.

La transformation de fonction de conversion `cast` appliquée est reflétée sur le canevas.

Types de données pris en charge pour les transformations de conversion

Metadata Insights prend en charge les types de données mentionnés dans le tableau ci-dessous lors de l'application de transformations de conversion.

Tableau 2 : Types de données pris en charge pour les transformations de conversion

Type	Description
String ou Varchar	Chaîne de caractères de longueur variable d'une longueur maximale égale à 4 000
Varbinary	Chaîne binaire de longueur variable d'une longueur maximale nominale égale à 8 192
Char	Caractère Unicode unique
Booléen	Bit unique, ou valeur booléenne, avec deux valeurs possibles
Byte ou Tinyint	Huit bits signés, numériques, de type intégral

Type	Description
Short ou Smallint	16 bits signés, numériques, de type intégral
Integer ou Serial	32 bits signés, numériques, de type intégral. Le type Serial implique également une valeur non nulle et comporte une valeur incrémentielle automatique commençant à 1. Les types Serial ne sont pas automatiquement UNIQUES.
Long ou Bigint	64 bits signés, numériques, de type intégral
Biginteger	Valeur numérique, de type intégral, à une précision arbitraire jusqu'à 1 000 chiffres
Float ou Real	32 bits, numériques, de type à virgule flottante, conformes à la norme IEEE 754
Double	64 bits signés, numériques, de type à virgule flottante, conformes à la norme IEEE 754
Bigdecimal ou Decimal	Valeur numérique, de type à virgule flottante, à une précision arbitraire jusqu'à 1 000 chiffres
Date	Datetime, représentant un seul jour (année, mois, jour)
Durée	Datetime, représentant une seule heure (heures, minutes, secondes, millisecondes)
Timestamp	Datetime, représentant une seule date et une seule heure (année, mois, jour, heures, minutes, secondes, millisecondes, nanosecondes)

Modification d'une transformation


Pour les transformations de colonne, de constante et de fonction, la modification peut être effectuée en cliquant sur l'icône Transformation du volet **Transformations** pour l'entité en question. En outre, en cas de transformations de constante et de fonction, vous pouvez double-cliquer sur le nœud Transformation sur le canevas et modifier les valeurs de la fenêtre contextuelle **Transformation de colonne '<nom de colonne>'** qui s'affiche.

1. Pour modifier une transformation, sélectionnez l'entité associée, puis cliquez sur l'onglet réductible **Transformations** sur le côté droit de la page.

Remarque : En cas de transformations de constante et de fonction, vous pouvez double-cliquer sur le nœud Transformation sur le canevas, qui vous permet d'afficher la

fenêtre contextuelle **Transformation de colonne** '<nom de colonne>' pour la modification.

Le volet **Transformations** s'affiche, indiquant le nom de tous les attributs de l'entité et le type de transformation appliqué. Au cas où aucune transformation n'est appliquée à l'attribut, la colonne **Transformation** reste vide.

2. Sélectionnez l'attribut auquel vous souhaitez appliquer la valeur transformée, puis cliquez sur l'icône **Transformation** .

En cas de transformations de constante et de fonction, vous pouvez également double-cliquer sur le nœud de transformation du canevas pour le modifier.


La fenêtre contextuelle **Transformation de colonne** '<nom de colonne>' apparaît.

3. Dans la fenêtre contextuelle, sélectionnez le paramètre de colonne, de constante ou de fonction que vous souhaitez modifier.
4. Sélectionnez une option dans **Colonne**, **Constante** ou **Fonction** pour définir le type de transformation à appliquer sur les valeurs des tables de Physical Models.

Option	Description
Colonne	Dérivez la valeur d'une colonne de table de Physical Models. Pour plus d'informations, reportez-vous à la section Application d'une transformation de colonne à la page 86.
Constante	Définissez la valeur obtenue comme constante. Pour plus d'informations, reportez-vous à la section Application d'une transformation de constante à la page 87.
Fonction	Définissez la valeur obtenue à l'aide des fonctions String, Numeric, Conversion et Datetime. Pour plus d'informations, reportez-vous à la section Application d'une transformation de fonction à la page 87.

La transformation modifiée est affichée sur le canevas pour l'entité logique correspondante.

Suppression d'une transformation

1. Supprimez une transformation en suivant l'une des étapes ci-dessous :
 - Sélectionnez le nœud de fonction de transformation ou le lien de transformation sur le canevas et appuyez sur la touche **Suppr** de votre clavier.
 - Dans le volet **Transformations** de l'entité logique, cliquez sur  en regard de la transformation que vous souhaitez supprimer.
2. Cliquez sur **Enregistrer**.

Critères

Utilisez le Générateur de critères pour définir un critère de filtrage des données extraites des colonnes Physical Model. Les données filtrées sont chargées dans les colonnes Logical Model.

Remarque : Cette étape est facultative.

Composants d'un critère

Le Générateur de critères d'une entité logique est constitué d'un groupe parent, qui comprend une hiérarchie de critères enfants et de groupes enfants.

Groupe Un groupe est défini à l'aide de plusieurs critères et groupes enfants. Vous pouvez créer une hiérarchie de groupes enfants et de critères jusqu'au niveau de votre choix au sein du groupe parent.

Pour chaque groupe, vous devez indiquer si les données extraites des tables Physical Model sources doivent remplir tous les critères, au moins un critère ou ne remplir aucun des critères définis de ce groupe.

Criterion Un critère est défini à l'aide d'un ou de deux opérandes et d'un opérateur logique.

L'opérateur logique sélectionné compare l'opérande gauche aux valeurs définies dans l'opérande droit ou place une condition sur l'opérande gauche.

Left Operand Colonne de la table Physical Model sur laquelle vous souhaitez définir le critère.

Right Operand L'opérande droit est défini comme l'un des élément suivants :

- Une constante
- Un ensemble de constantes
- Une colonne de l'une des tables Physical Model
- Un ensemble de commandes de l'une des tables Physical Model

Opérateur L'opérateur logique définit la condition à remplir pour le critère. L'opérateur peut effectuer les opérations suivantes :

- Placer une condition sur l'opérande gauche.
- Placer une condition comparant l'opérande gauche à l'opérande droit.


Création d'un critère



La section **Critères** s'affiche pour l'entité Logical Model sélectionnée uniquement si au moins une transformation lui est définie.

1. Sur la page **Créer mappage <nom de Logical Model>**, sélectionnez l'entité logique requise, puis cliquez sur l'icône **Générateur de critères** de la barre d'outils .

Remarque : Le volet **Transformation** n'affiche pas les détails de la table Physical Model. Si vous cliquez sur une table Physical Model après une entité et si vous ouvrez le volet **Transformation**, il affiche les détails de la dernière entité sélectionnée.

La fenêtre contextuelle **Générateur de critères de la table logique <nom de la table>** s'affiche avec la liste déroulante **Groupe répondant au critère**.

Remarque : Pour ajouter un autre groupe, passez la souris sur le groupe et cliquez sur l'icône **Ajouter un groupe**  qui s'affiche.

2. Sélectionnez une méthode de regroupement des critères en cliquant sur la liste déroulante **Regrouper les critères satisfaits**. Indiquez s'il est nécessaire de remplir tous les critères, au moins un critère ou aucun des critères définis dans le groupe de critères.
3. Pour ajouter un groupe enfant au groupe parent, passez la souris sur le groupe parent et cliquez sur l'icône **Ajouter un groupe**  qui s'affiche.
4. Pour ajouter un critère à un groupe, passez la souris sur le groupe parent et cliquez sur l'icône **Ajouter une comparaison**  qui s'affiche. Le critère ajouté s'affiche sous le groupe.
 - a) Pour définir l'opérande gauche du critère, dans **Expression Builder**, cliquez sur **<Colonne 1>** pour sélectionner l'une des colonnes de table physique.
 - a) Pour sélectionner un opérateur logique dans la liste fournie, dans **Expression Builder**, cliquez sur **<comparaison>**.
Si l'opérateur logique sélectionné requiert un opérande droit, un paramètre fictif à définir est affiché pour l'opérande droit.
 - b) Si l'opérande droit est nécessaire, cliquez sur **<Colonne 2>**.
L'opérande droit peut être défini à l'aide d'une constante ou d'une colonne de table Physical Model.

Constante	Saisissez la constante, puis cliquez sur Ajouter constante .
colonne(s)	Sélectionnez l'une des colonnes Physical Model disponibles.


Suivant le nombre d'opérandes attendu par l'opérateur logique dans l'opérande droit, définissez l'opérande droit comme suit :

- Une constante
- Un ensemble de constantes
- Une colonne de l'une des tables Physical Model
- Un ensemble de commandes de l'une des tables Physical Model



5. Répétez les étapes 3 et 4 autant de fois que nécessaire pour créer les critères souhaités.
6. Cliquez sur **OK**.

L'ensemble, l'un ou aucun des critères enfants définis doit être défini sur true pour permettre l'extraction des données requises des tables Physical Model connectées dans les colonnes Logical Model cibles. Ceci est défini comme pour le critère parent.

Modification d'un critère

1. Sélectionnez l'entité logique requise, puis cliquez sur l'icône **Générateur de critères** de la barre d'outils .
La fenêtre contextuelle **Générateur de critères pour la table logique <nom de la table>** apparaît avec les groupes et les critères définis.
2. Sélectionnez le groupe requis, un groupe enfant ou un critère.
3. Apportez les modifications nécessaires.
4. Cliquez sur **OK**.

Suppression d'un critère

1. Sélectionnez l'entité logique requise, puis cliquez sur l'icône **Générateur de critères** de la barre d'outils .
La fenêtre contextuelle **Générateur de critères pour la table logique <nom de la table>** apparaît avec les groupes et les critères définis.
2. Sélectionnez le groupe requis, un groupe enfant ou un critère.
3. Cliquez sur l'icône **Supprimer** .
4. Cliquez sur **Enregistrer**.

Jointures

Utilisez le **Générateur de jointure** pour combiner les données des colonnes de deux tables Physical Model avant de les fournir aux colonnes Logical Model.

Exemple : La colonne `Name` de la table `Product` de votre Logical Model est associée à la colonne `Product Name` de la table physique `Product_A`. Cependant, vous souhaitez extraire les données d'une autre table physique `Product_B`. Dans ce cas, vous pouvez joindre la colonne `Product Name` de la table physique `Product_A` et la colonne `Product Name` de la table physique `Product_B` de sorte que les données fournies au Logical Model soient une combinaison des données entre les deux tables physiques (en fonction du type de jointure sélectionné).

Remarque : Vous devez définir une ou plusieurs transformations sur la table logique avant d'utiliser le Générateur de jointure.


Remarque : Vous ne pouvez effectuer des jointures que si vous disposez d'au moins deux tables Physical Model glissées sur le canevas.

Création d'une jointure

1. Sur la page **Créer mappage <nom de Logical Model>**, sélectionnez l'entité logique requise, puis cliquez sur l'icône **Générateur de jointure** de la barre d'outils .

Remarque : Le volet **Transformation** n'affiche pas les détails de la table Physical Model. Si vous cliquez sur une table Physical Model après une entité et si vous ouvrez le volet **Transformation**, il affiche les détails de la dernière entité sélectionnée.

La fenêtre contextuelle **Générateur de jointure pour la table logique <nom de la table>** s'affiche.

2. Cliquez sur le bouton **Ajouter une table**.
Une table s'affiche avec les colonnes **Table gauche**, **Table droite** et **Type de jointure**.
3. Utilisez les options de liste déroulante **<sélectionner une table>** et **<sélectionner une colonne>** des colonnes **Table gauche** et **Table droite** pour spécifier les colonnes (des tables Physical Model) dont les données doivent être jointes.
4. Dans la liste déroulante **Type de jointure**, sélectionnez le type de jointure que vous souhaitez entre les colonnes. Les options sont :
 - **Jointure interne** : renvoie des enregistrements de correspondance entre les colonnes des tables gauche et droite sélectionnées.
 - **Jointure gauche** : renvoie tous les enregistrements de la colonne gauche, en plus de tous les enregistrements de correspondance de la colonne droite.
 - **Jointure droite** : renvoie tous les enregistrements de la colonne droite, en plus de tous les enregistrements de correspondance de la colonne gauche.
 - **Jointure externe complète** : renvoie tous les enregistrements des colonnes gauche et droite.
5. Pour établir des jointures sur plusieurs colonnes de la table, placez le curseur sur la ligne de la table, puis cliquez sur l'icône **Ajouter une comparaison**  qui s'affiche.
Une nouvelle ligne **<sélectionner une colonne>** s'affiche.
6. Répétez les étapes 4 et 5 pour définir les critères de jointure.
7. Pour établir la jointure entre un autre ensemble de tables, cliquez sur le bouton **Ajouter table**.

Remarque : Vous devez faire glisser l'ensemble de tables jusqu'au canevas pour pouvoir établir une jointure entre elles.

Un nouvel ensemble de lignes **<sélectionner une table>** et **<sélectionner une colonne>** s'affichent.

8. Répétez les étapes 3 et 4 pour définir la jointure.
9. Cliquez sur **OK**.

La jointure entre les tables est établie et les données sont fournies dans la colonne Logical Model à partir des colonnes Physical Model sélectionnées en fonction de la jointure définie.

Remarque : Les jointures ne sont représentées par aucun lien physique entre les tables sur le canevas.

Modification d'une jointure

1. Sur la page **Créer mappage <nom de Logical Model>**, sélectionnez l'entité logique requise, puis cliquez sur l'icône **Générateur de jointure** de la barre d'outils .

Remarque : Le volet **Transformation** n'affiche pas les détails de la table Physical Model. Si vous cliquez sur une table Physical Model après une entité et si vous ouvrez le volet **Transformation**, il affiche les détails de la dernière entité sélectionnée.

La fenêtre contextuelle **Générateur de jointure pour la table logique <nom de la table>** s'affiche.

2. Modifiez votre sélection, selon vos besoins, puis cliquez sur le bouton **OK**.

Remarque : Pour plus d'informations sur l'utilisation des champs pour la modification, reportez-vous à la section **Création d'une jointure** à la page 97.

Les définitions de jointure sont modifiées pour les tables et les colonnes sélectionnées.


Remarque : Les jointures ne sont représentées par aucun lien physique entre les tables sur le canevas.

Suppression d'une jointure

1. Sur la page **Créer mappage <nom de Logical Model>**, sélectionnez l'entité logique requise, puis cliquez sur l'icône **Générateur de jointure** de la barre d'outils .

Remarque : Le volet **Transformation** n'affiche pas les détails de la table Physical Model. Si vous cliquez sur une table Physical Model après une entité et si vous ouvrez le volet **Transformation**, il affiche les détails de la dernière entité sélectionnée.

La fenêtre contextuelle **Générateur de jointure pour la table logique <nom de la table>** s'affiche.

2. Pour supprimer une colonne ou une ligne de la table, passez le curseur sur la ligne, puis cliquez sur l'icône Supprimer .
3. Cliquez sur le bouton **OK**.

Model Store

Un Model Store est une représentation virtuelle des Logical Models et des Physical Models mappés. Une fois déployé, il devient accessible à partir de Spectrum™ Technology Platform et d'autres applications client.

Les Logical Models et les Physical Models mappés d'un Model Store comportent des données fédérées de plusieurs sources, telles que des bases de données, des serveurs de fichiers et des services Cloud. Vous pouvez concevoir et personnaliser votre Model Store en fonction de vos besoins et, pour chacun des Logical Models du Model Store, vous pouvez définir la configuration du cache, ce qui garantit l'actualisation périodique des données. Toutes ces fonctionnalités font d'un Model Store une option pratique pour stocker et utiliser des données personnalisées sans avoir à subir des délais prolongés dus à la congestion du réseau. Ce cache local améliore les performances des requêtes lorsque des flux ou des services de Spectrum™ Technology Platform référencent ces données.

Vous pouvez créer plusieurs Model Stores en fonction de vos différents besoins commerciaux et les déployer. Cependant, pour accéder à un Model Store déployé via Spectrum™ Technology Platform, vous devez d'abord créer une connexion vers lui à l'aide de **Management Console**. Pour plus d'informations, reportez-vous à la section [Connexion à un Model Store](#) à la page 51.

Vous pouvez également modifier et supprimer un Model Store ; cependant, pour ce faire, vous devez commencer par annuler son déploiement.


Création d'un Model Store

Pour créer un Model Store, vous devez effectuer ces tâches sur la page **Créer un Model Store**.

1. Dans un navigateur Web, accédez à :

`http://server:port/metadata-insights`

Où : *server* est l'adresse IP ou le nom de serveur de votre serveur Spectrum™ Technology Platform et *port* le port HTTP. Par défaut, le port HTTP est 8080.

2. Connectez-vous à l'aide de vos informations d'identification.
3. Cliquez sur **Modélisation > Model Store**.
4. Cliquez sur l'icône **Créer un Model Store** . La page **Créer un Model Store** s'affiche, répertoriant le **Nom** et le **Type** (physique ou logique) de tous les modèles disponibles.
5. Dans la zone de texte **Nom**, saisissez un nom pour le Model Store.
6. Cochez la case correspondant aux modèles que vous souhaitez inclure dans le Model Store. Pour sélectionner tous les modèles répertoriés, cochez la case **Inclure** dans la ligne d'en-tête. Pour supprimer un modèle sélectionné, décochez la ligne correspondante.

Remarque : Vous pouvez utiliser la zone de texte à saisie anticipée **Filtre** pour personnaliser l'affichage de la grille. Par exemple, si les modèles que vous souhaitez sélectionner commencent par le texte « Cust », vous pouvez saisir `Cust` dans la zone de filtre, et tous les modèles dont le nom comporte « Cust » s'affichent.

Les modèles sélectionnés sont inclus dans le Model Store.

Remarque : Si vous cliquez sur le bouton **Enregistrer** sans inclure aucun modèle, un message d'erreur vous demande d'inclure au moins un modèle de la table.

7. Pour définir la configuration du cache des Logical Models inclus dans le Model Store, cliquez sur le bouton **Configuration du cache** respectif.

Remarque : Le bouton **Configuration du cache** s'affiche une fois que vous cochez la case **Inclure** du Logical Model.

La fenêtre contextuelle **Configuration du cache pour le modèle <nom du modèle>** s'affiche. Pour plus d'informations sur la configuration du cache, reportez-vous à la section **Configuration du cache** à la page 108.

Remarque : Vous pouvez configurer le cache d'un seul Logical Model à la fois.

8. Pour enregistrer et déployer simultanément le Model Store, cliquez sur **Enregistrer et déployer**. Si vous avez sélectionné un Logical Model, mais que vous n'avez pas sélectionné tous les Physical Models mappés, un message vous invite à ajouter également les Physical Models dépendants. Cliquez sur **Oui** et continuez.

Pour déployer correctement un Model Store, vous devez éviter ces deux scénarios.

- Un Logical Model constituant n'est mappé vers aucun Physical Model.
- L'un des Physical Models mappés d'un Logical Model n'a pas été sélectionné pour être ajouté.

Le Model Store est enregistré et déployé pour être utilisé dans l'application client.

Remarque : Si vous n'avez pas respecté l'un des critères susmentionnés, votre Model Store est tout de même enregistré et un message d'erreur s'affiche, indiquant que le Model Store n'a pas été déployé. Pour plus d'informations, reportez-vous à la section **Déploiement d'un Model Store** à la page 101.

9. Cliquez sur **Enregistrer** pour enregistrer le Model Store et le déployer ultérieurement.


Un Model Store doit être déployé pour être accessible depuis Spectrum™ Technology Platform ainsi que depuis d'autres applications client.

- Pour déployer un Model Store, reportez-vous à la section **Déploiement d'un Model Store** à la page 101.
- Pour accéder à un Model Store depuis Spectrum™ Technology Platform, vous devez d'abord créer une connexion vers lui à l'aide de **Management Console**. Pour plus d'informations, reportez-vous à la section **Connexion à un Model Store** à la page 51 du **Guide d'administration**.
- Pour accéder à un Model Store depuis une application client autre que Spectrum™ Technology Platform, vous devez télécharger le pilote JDBC et l'installer manuellement. Pour plus d'informations sur le téléchargement et l'installation du pilote JDBC, reportez-vous à la section **Installation d'un pilote JDBC** à la page 103.

Modification d'un Model Store

Pour modifier un Model Store existant, vous devez effectuer ces tâches sur la page **Modifier un Model Store**.

Seul un Model Store non déployé peut être modifié. Pour modifier un Model Store déployé, vous devez tout d'abord annuler son déploiement.

1. Sur la page d'accueil **Metadata Insights**, cliquez sur **Modélisation > Model Store**.
2. Dans la liste de Model Stores affichée, sélectionnez le Model Store que vous souhaitez modifier et cliquez sur l'icône **Modifier un Model Store** .

Remarque : Vous ne pouvez modifier que les Model Stores qui ne sont pas encore déployés. Pour un Model Store déployé, l'icône **Modifier un Model Store** reste indisponible tant que vous n'annulez pas son déploiement.

La page **Modifier un Model Store** s'affiche.

3. Ajoutez ou supprimez les Physical Models et les Logical Models, si nécessaire, en cliquant sur les lignes **Inclure** correspondantes.
4. Cliquez sur **Enregistrer** ou sur **Enregistrer et déployer**, le cas échéant.
Pour déployer correctement un Model Store, vous devez éviter ces deux scénarios.
 - Un Logical Model constituant n'est mappé vers aucun Physical Model.
 - L'un des Physical Models mappés d'un Logical Model n'a pas été sélectionné pour être ajouté.

Suppression d'un Model Store

Seul un Model Store non déployé peut être supprimé. Pour supprimer un Model Store déployé, vous devez tout d'abord annuler son déploiement.


1. Sur la page d'accueil **Metadata Insights**, cliquez sur **Modélisation > Model Store**.
2. Dans la liste de modèles affichée, sélectionnez le Model Store que vous souhaitez supprimer et cliquez sur l'icône **Supprimer**.

Un message confirme la suppression du Model Store sélectionné.


Déploiement d'un Model Store

Vous pouvez déployer un Model Store disponible pour l'utiliser dans votre application, ou annuler le déploiement d'un Model Store déjà déployé, de sorte qu'il ne puisse plus être utilisé. Pour ce faire, effectuez ces tâches sur la page **Model Store**.

1. Sur la page d'accueil **Metadata Insights**, cliquez sur **Modélisation > Model Store**.
2. Dans la liste de modèles affichée, sélectionnez le modèle que vous souhaitez déployer (ou dont vous souhaitez annuler le déploiement) et cliquez sur l'une de ces icônes, selon les besoins.

- **Déployer un Model Store**  : pour déployer le Model Store sélectionné. Cette icône est activée uniquement lorsque vous sélectionnez un Model Store non déployé dans la liste.

Remarque : Le déploiement échoue si le Model Store contient un Logical Model dont toutes les entités sont non mappées.

- **Annuler le déploiement d'un Model Store**  : pour annuler le déploiement d'un Model Store déjà déployé. Cette icône est activée uniquement lorsque vous sélectionnez un Model Store déployé dans la liste.


Remarque : Vous ne pouvez déployer qu'un seul Model Store à la fois, de même que vous ne pouvez annuler le déploiement que d'un seul Model Store à la fois. Si vous sélectionnez plusieurs Model Stores, les icônes se désactivent.

Remarque : Si un Physical Model ou un Logical Model dépendant est supprimé d'un Model Store, le Model Store devient inutilisable.

Prévisualisation d'un Model Store

Pour afficher un aperçu d'un Model Store déployé, sélectionnez-le dans la liste des Model Stores, cliquez sur **Aperçu** et exécutez la requête SQL nécessaire. Le résultat de l'aperçu affiche un maximum de 50 enregistrements.

Remarque : Il n'est pas possible d'afficher un aperçu d'un Model Store non déployé.

1. Sur la page d'accueil **Metadata Insights**, cliquez sur **Modélisation > Model Store**.
2. Dans la liste de Model Stores affichée, sélectionnez le Model Store dont vous souhaitez afficher un aperçu et cliquez sur l'icône **Aperçu** .

La page **Afficher un aperçu d'un Model Store - <nom du modèle>** apparaît.

3. Dans la zone de texte **Requête SQL**, saisissez la requête permettant de récupérer les données de votre choix du Model Store sélectionné. Encadrez le nom de table de guillemets doubles.

Sélectionnez `* from "EMPLOYEE_SALARY" ;`

Remarque : Si le même nom de table existe déjà dans deux Physical Models différents, vous devez fournir le nom `<physicalmodelname>.table` pour que la requête soit correctement exécutée.

4. Cliquez sur **Aperçu**.

Un aperçu des données extraites des Physical Models et des Logical Models constituants s'affiche.

Accès à un Model Store

Un Model Store déployé est accessible depuis Spectrum™ Technology Platform ainsi que depuis d'autres applications client. Si vous y accédez depuis Spectrum™ Technology Platform, vous devez uniquement créer une connexion au Model Store requis à l'aide de **Management Console** lors de l'installation du pilote JDBC lors de l'installation de Spectrum™ Technology Platform. Pour plus d'informations sur la création d'une connexion à un Model Store, reportez-vous à la section **Connexion à un Model Store** à la page 51 du **Guide d'administration**.

En revanche, si vous y accédez à partir de toute autre application, vous devez télécharger le pilote JDBC et l'installer manuellement. Une fois installé, vous devez créer une connexion à Metadata Insights à l'aide d'un code simple, pour pouvoir utiliser le Model Store.

Installation d'un pilote JDBC

Procédez comme suit pour télécharger le pilote JDBC, installez-le et créer la connexion au Model Store.

1. Sur la page d'accueil Spectrum™ Technology Platform, cliquez sur **Outils client de la plateforme**. Dans les options du volet droit, développez le menu en cascade **Pilote**, puis cliquez sur le bouton **Télécharger le fichier Jar** du pilote JDBC.
2. Enregistrez le fichier *modelstore-jdbc.jar* dans le répertoire de navigateur par défaut.

Remarque : Vous n'avez pas besoin de redémarrer le serveur.

Une fois le pilote JDBC installé, utilisez les détails suivants pour vous connecter au serveur Spectrum™ via un simple code ou n'importe quelle application :

1. Nom d'utilisateur Spectrum™
2. Mot de passe pour authentifier le nom d'utilisateur Spectrum™
3. JDBC Driver Class Name =
`com.pb.spectrum.modelstore.driver.modelstoreDriver`
4. Connection String Template =
`jdbc:spectrum://${host}:${port}/${modelstore}`.

Saisissez l'adresse IP du serveur ou le nom du poste à la place de `{host}`, le port de Metadata Insights dans `{port}` et le nom du Model Store auquel vous souhaitez accéder dans `{modelstore}`.

Remarque : La valeur par défaut de `{port}` est 32750.

Vous pouvez maintenant vous connecter à n'importe quel Model Store de votre choix.

Installation d'un pilote ODBC

Vous pouvez accéder à un Model Store via une connexion ODBC dans les applications prenant en charge une interface ODBC, comme MS Excel et MS Access.

1. La version du pilote ODBC dépend de l'architecture de l'application client.
 - Si la version de l'application client est de 32 bits, téléchargez le pilote PostgreSQL ODBC de 32 bits [ici](#).
 - Si la version de l'application client est de 64 bits, téléchargez le pilote PostgreSQL ODBC de 64 bits [ici](#).

Par exemple, si vous utilisez la version de 32 bits de MS Office, téléchargez le pilote ODBC de 32 bits.

2. Exécutez le programme d'installation `psqlodbc.msi` s'il s'agit d'un pilote ODBC de 32 bits. S'il s'agit d'un pilote ODBC de 64 bits, exécutez le programme d'installation `.exe`.

Suivez les étapes pour terminer l'installation du pilote.

Le pilote ODBC est installé.

Création d'un Data Source Name (DSN – nom de source de données)

Ces instructions sont spécifiques aux plates-formes Windows.

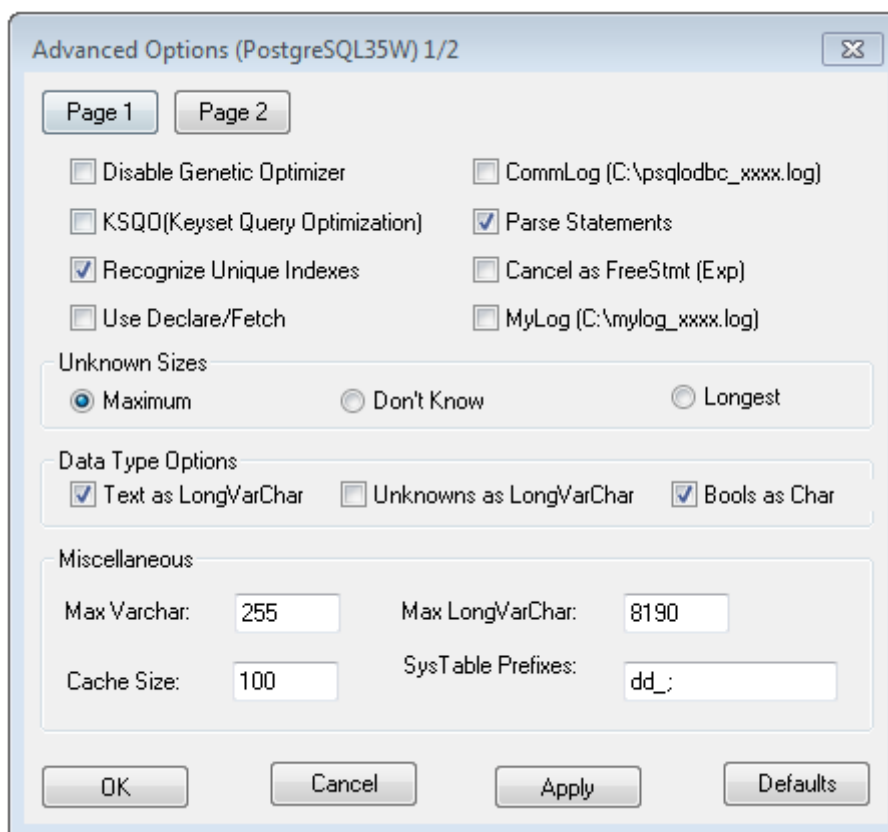
Avertissement : Assurez-vous que vous êtes connecté à l'ordinateur client avec des *droits administratifs*.

1. Dans le menu **Démarrer**, accédez à **Panneau de configuration > Outils d'administration**.
2. Double-cliquez sur **Sources de données (ODBC)**.
La fenêtre **Administrateur de sources de données ODBC** s'ouvre.
3. Cliquez sur l'onglet **DSN utilisateur** ou **DSN système**, suivant le type de DSN que vous souhaitez créer.
4. Cliquez sur **Ajouter**.
5. Dans la liste des pilotes installés sur l'ordinateur, sélectionnez `PostgreSQL ANSI`, puis cliquez sur **Terminer**.
La fenêtre contextuelle **Configuration du lecteur PostgreSQL ANSI ODBC (psqlODBC)** s'ouvre.

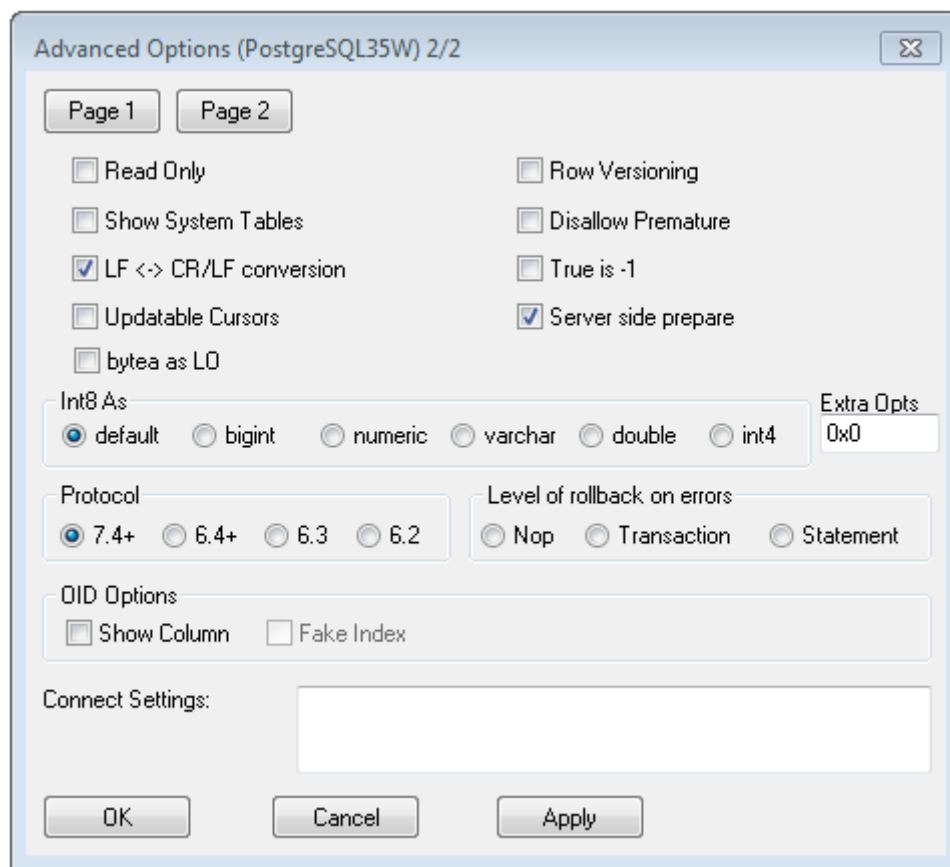
Avertissement : Sélectionnez `PostgreSQL ANSI` ou `PostgreSQL Unicode`, suivant l'application client via laquelle la connexion ODBC doit être utilisée.

6. Dans la fenêtre contextuelle, saisissez les détails suivants pour mapper ce DSN au Model Store: de votre choix :
 - a) Dans le champ **Source de données**, saisissez le nom à attribuer au DSN.
 - b) Dans le champ **Base de données**, saisissez le nom du Model Store auquel vous souhaitez vous connecter via ce DSN.

- c) Dans le champ **Serveur**, saisissez le nom de l'ordinateur sur lequel le serveur Spectrum™ Technology Platform est exécuté.
- d) Dans le champ **Port**, saisissez la valeur de la propriété `edf.odbc.socket.port`, telle que définie dans le fichier `federation_config.properties`. Reportez-vous à la section **Propriétés de configuration communes**.
Par défaut, le port est 32751.
- e) Dans le champ **Nom d'utilisateur**, saisissez le nom d'utilisateur du serveur Spectrum™ Technology Platform.
- f) Dans le champ **Mot de passe**, saisissez le mot de passe du serveur Spectrum™ Technology Platform.
- g) Cliquez sur **Source de données** dans la section **Options**.
La fenêtre contextuelle **Options avancées (PostgreSQL30)** s'ouvre.
- h) Sur la **Page 1**, configurez les options comme suit :



- i) Cliquez sur la **Page 2** et configurez les options comme suit :



- j) Cliquez sur **OK** pour fermer la fenêtre contextuelle **Options avancées (PostgreSQL30)**.
- k) Cliquez sur **Test** pour vérifier que le DSN créé se connecte correctement au Model Store configuré.

Important : Assurez-vous que le Model Store est déployé pour tester correctement le DSN.

- l) Cliquez sur **Enregistrer**.

Le DSN créé vous permet de vous connecter au Model Store configuré à l'aide d'applications telles que MS Excel et MS Access.

Accès à un DSN via MS Excel

Vous pouvez vous connecter à un DSN ODBC configuré depuis une application client et accéder au Model Store mappé pour effectuer les opérations nécessaires. Un DSN est accessible par différentes applications client telles que MS Excel et MS Access. Cet exemple illustre les étapes à suivre pour se connecter à un DSN ODBC et accéder à un Model Store dans MS Excel.

1. Ouvrez MS Excel.
2. Accédez à **Données > Autres sources > Assistant Connexion de données**. L'**Assistant Connexion de données** s'ouvre.
3. Sélectionnez **DSN ODBC** comme type de source de données.
4. Cliquez sur **Suivant**.

5. Dans la liste de sources de données ODBC, sélectionnez le DSN auquel vous souhaitez vous connecter.
6. Cliquez sur **Suivant**.
7. Assurez-vous que le Model Store correct est sélectionné dans la liste déroulante de bases de données.
8. Si vous souhaitez sélectionner l'une des tables répertoriées du Model Store, cochez la case **Se connecter à une table spécifique** et sélectionnez la table de votre choix.
Sinon, pour sélectionner toutes les tables répertoriées, laissez cette case décochée.
9. Cliquez sur **Suivant**.

Le contenu de la ou des tables sélectionnées est importé et affiché dans la feuille de calcul.

Remarque : Les données affichées dans MS Excel sont en lecture seule.

Cependant, il est possible d'y écrire lors de l'accès à un DSN à l'aide de MS Access et d'autres applications client.

Remarque : L'exemple illustre la procédure pour un pilote ODBC de 32 bits.

Configuration du cache

Vous pouvez mettre en cache les données de tous les Logical Models (et de leurs entités) dans un Model Store. Cela réduit les délais lors de l'accès aux données dus à la congestion du réseau.

La mise en cache peut être activée ou désactivée pour toutes les entités ou pour les entités sélectionnées d'un Logical Model. Vous pouvez définir la durée pendant laquelle les données mises en cache restent accessibles, suite à quoi le cache expire et les données doivent être rechargées.

Chargement du cache Le cache de chaque modèle est chargé lorsque la première requête est exécutée pour y accéder une fois la mise en cache activée.

Une fois la mise en cache activée pour un Model Store, l'exécution de la première requête dessus peut prendre plus de temps que d'habitude. Cela est dû au fait que les données mises en cache doivent d'abord être actualisées depuis la source de données, suite à quoi les données requises sont récupérées du cache actualisé.

Durée de vie du modèle Il s'agit de la durée entre le chargement d'un cache et sa purge suivante. Une fois qu'un cache expire, les données du cache sont rechargées à partir des sources de données lors de l'exécution de la requête immédiate suivante.

La **Durée de vie (TTL)** suivante est calculée à partir du moment où le rechargement des données en cache est terminé. Une fois le délai d'expiration atteint, les données du cache sont de nouveau purgées.

La **Durée de vie (TTL)** par défaut est de 120 minutes.


Configuration du cache

Vous pouvez configurer la durée du cache des Logical Models d'un Model Store à l'aide du bouton **Configuration du cache**.

1. Dans un navigateur Web, accédez à :

`http://server:port/metadata-insights`

Où : *server* est l'adresse IP ou le nom de serveur de votre serveur Spectrum™ Technology Platform et *port* le port HTTP. Par défaut, le port HTTP est 8080.

2. Connectez-vous à l'aide de vos informations d'identification.
3. Sur la page d'accueil **Metadata Insights**, cliquez sur **Modélisation > Model Store**.
4. Dans la liste de Model Stores affichée, sélectionnez le Model Store dont vous souhaitez configurer le cache, puis cliquez sur l'icône **Modifier un Model Store** .

Remarque : Vous ne pouvez pas définir la configuration du cache de Model Stores déjà déployés. Dans ce cas, vous devez annuler le déploiement du Model Store. L'icône **Modifier un Model Store** est activée après le déploiement.

La page **Modifier un Model Store** s'affiche, répertoriant tous les modèles de ce Model Store.

5. Cochez la case **Inclure** du Logical Model et cliquez sur le bouton **Configuration du cache** qui s'affiche.

Remarque : Ce bouton s'affiche uniquement pour les Logical Models du Model Store.

Une fois que la configuration est définie pour un Logical Model du Model Store, le bouton **Configuration du cache** reste activé tant que vous n'avez pas décoché la case **Inclure**.

La fenêtre contextuelle **Configuration du cache pour le modèle <nom du Logical Model>** apparaît, affichant une durée de cache par défaut de **120** minutes (dans le champ **Durée de vie (TTL) du modèle**) pour toutes les entités du modèle.

6. Dans le champ **Durée de vie (TTL) du modèle**, saisissez la durée (en minutes) pendant laquelle vous souhaitez activer la mise en cache. La durée par défaut est de 120 minutes.

Remarque : La durée que vous saisissez dans ce champ est automatiquement renseignée dans les champs **Durée de vie (TTL)** de toutes les entités du modèle. Cependant, vous pouvez remplacer la durée du cache pour toute entité en activant la sélection **Remplacer TTL**.

7. Pour activer la mise en cache des entités du modèle, cochez la case **Activer le cache** respective. Pour activer la mise en cache de toutes les entités du modèle, cochez la case **Activer le cache** de l'en-tête.
8. Pour personnaliser la durée du cache pour l'une des entités du modèle, activez le curseur **Remplacer TTL** et saisissez la durée de remplacement du cache requise dans le champ **Durée de vie (TTL)** respectif.

9. Pour effacer tous les remplacements de configuration du cache du modèle, cliquez sur le bouton **Effacer tous les remplacements**.

Les personnalisations effectuées dans les champs **Remplacer TTL** et **Durée de vie (TTL)** sont remplacées et la **Durée de vie (TTL)** de toutes les entités est définie sur la valeur définie dans le champ **Durée de vie (TTL) du modèle**.

10. Cliquez sur le bouton **OK**.

La durée du cache est définie, telle que spécifiée pour toutes les entités du Logical Model, et vous revenez à la page **Modifier un Model Store**.

3 - Profilage

La prise de décisions commerciales éclairées nécessite des données de qualité. C'est pourquoi il est important pour vous d'avoir confiance dans l'exhaustivité, l'exactitude et la validité de vos données. Les enregistrements incomplets, les champs non conformes et le manque de contexte peuvent entraîner la fourniture de données trompeuses ou inappropriées à vos utilisateurs commerciaux, ce qui peut aboutir à des décisions faussées.

Le profilage des données peut vous donner confiance dans vos données. Le profilage analyse vos données et génère des rapports qui identifient les problèmes en matière d'exactitude, d'exhaustivité et de validité. Grâce à ces rapports, vous pouvez entreprendre des actions pour corriger les données incorrectes ou non conformes.

Metadata Insights fournit des outils de profilage qui permettent d'exécuter un profilage sur vos actifs de données, ainsi que sur les données qui alimentent les Logical Models et les Physical Models définis dans Metadata Insights. Grâce à ces informations, vous pouvez déterminer la fiabilité de vos données, créer des règles de qualité des données et appliquer des routines de standardisation et de normalisation pour résoudre les problèmes de qualité des données.

Les tâches de profilage impliquées sont les suivantes :

- Création d'un profil pour vos sources de données, qui implique :
 - Sélection des sources de données
 - Configuration des règles, de la taille d'échantillon et des notifications
 - Définition du profil
- Analyse d'un profil et création d'un calendrier pour l'analyse, si nécessaire
- Affichage de rapports d'analyse

Remarque : Pour pouvoir utiliser la fonction de profilage de Metadata Insights, vous devez avoir installé le module Data Integration.

In this section

Création d'un profil	112
Spécification des paramètres de profilage par défaut	123
Analyse d'un profil	124
Affichage de rapports d'analyse du profil	129

Création d'un profil

Pour pouvoir exécuter les règles d'analyse de la qualité des données, vous devez créer un profil. Un profil peut inclure le jeu de données d'une de ces sources :

- Fichiers plats
- Logical Models ou Physical Models créés à l'aide de la fonction de modélisation de Metadata Insights
- Connexions configurées dans Spectrum™ Technology Platform

Dans Metadata Insights, la création d'un profil implique ces trois phases :

1. Ajout de sources de données : votre source de données peut être l'une des sources de données suivantes :
 - Fichiers plats de votre système ou de votre serveur
 - Logical Models ou Physical Models créés à l'aide de la fonction de modélisation de Metadata Insights
 - Connexions configurées dans Spectrum™ Technology Platform
2. Configuration des règles d'analyse des données, définition de la taille d'échantillon et configuration de la notification
3. Définition du profil


Ajout de données à partir de fichiers plats

Vous pouvez créer votre profil à partir d'enregistrements figurant dans des fichiers plats sur votre ordinateur ou sur le serveur. Pour ce faire, procédez comme suit :


1. Dans un navigateur Web, accédez à :

`http://server:port/metadata-insights`

Où : *server* est l'adresse IP ou le nom de serveur de votre serveur Spectrum™ Technology Platform et *port* le port HTTP. Par défaut, le port HTTP est 8080.

2. Connectez-vous à l'aide de vos informations d'identification.
3. Sur la page **Accueil**, cliquez sur **Profilage**.
4. Cliquez sur le bouton **Ajouter profil** . La page **Ajouter profil** s'affiche avec un assistant affichant les trois phases de création d'un profil : **Sélectionner la source** > **Configurer le profil** > **Terminer**.
5. Dans les options **Sélectionner une source de données**, sélectionnez l'option **Fichier plat**, puis cliquez sur le bouton **Ajouter des fichiers**.

La fenêtre contextuelle **Sélectionner un fichier** s'affiche.

6. Utilisez la zone de sélection du chemin d'accès aux fichiers  pour accéder au chemin d'accès sur votre poste local ou sur le serveur où se trouvent les fichiers. Une liste de fichiers s'affiche sous la zone de sélection du chemin d'accès aux fichiers. Vous pouvez filtrer les fichiers requis dans la liste à l'aide de la zone de texte à saisie anticipée **Filtre**.
7. Sélectionnez tous les fichiers dont vous avez besoin pour créer le profil et cliquez sur **OK**. Les fichiers sélectionnés sont affichés sur le côté gauche de la page, sous le bouton **Ajouter des fichiers**.
8. Pour ajouter d'autres fichiers, cliquez de nouveau sur le bouton **Ajouter des fichiers** et répétez les étapes 6 et 7.

Remarque : Vous pouvez ajouter des fichiers de différents dossiers figurant sur votre système local, sur votre serveur et sur le serveur FTP, dans n'importe quelle combinaison souhaitée. Pour utiliser le serveur FTP, vous devez le configurer à l'aide de Management Console. Pour obtenir des informations détaillées, reportez-vous au Guide d'administration.

9. Cliquez sur un fichier de la liste pour afficher ces détails :
 - Le nombre de fichiers sélectionnés. Par exemple : *7 fichier(s) sélectionnés*
 - Codage couleur pour les métadonnées du fichier
 - **Vert** : métadonnées automatiquement détectées
 - **Bleu** : métadonnées configurées par l'utilisateur
 - **Rouge** : métadonnées non détectées. Ces fichiers sont ignorés lors du profilage. Cependant, vous pouvez configurer les métadonnées de ces fichiers à l'aide des champs **Paramètres de fichiers**, comme décrit dans cette tâche ci-dessous. Après la configuration manuelle, ces fichiers deviennent bleus.
 - Chemin d'accès au fichier. Par exemple : *E:/Sample Files/profile/fileOne.txt*
 - Paramètres de fichiers
 - Aperçu des enregistrements du fichier
10. Pour modifier les paramètres de l'un des fichiers, sélectionnez-le, puis modifiez-le, tel que décrit dans ce tableau.

Paramètres	Description
Encodage des caractères	L'encodage du fichier texte. Sélectionnez l'un des éléments suivants :
UTF-8	Prend en charge tous les caractères Unicode et est rétrocompatible avec le format ASCII. Pour plus d'informations sur UTF, reportez-vous à unicode.org/faq/utf_bom.html .
UTF-16	Prend en charge tous les caractères Unicode, mais n'est pas rétrocompatible avec le format ASCII. Pour plus d'informations sur UTF, reportez-vous à unicode.org/faq/utf_bom.html .
US-ASCII	L'encodage des caractères en fonction de l'ordre alphabétique anglais.
UTF-16BE	Encodage UTF-16 avec sérialisation d'octets gros-boutiste (octet le plus significatif en premier).
UTF-16LE	Encodage UTF-16 avec sérialisation d'octets petit-boutiste (octet le moins significatif en premier).
ISO-8859-1	Chiffrement de caractères ASCII utilisé habituellement pour les langues européennes occidentales. Également désigné par l'acronyme Latin-1.
ISO-8859-3	Chiffrement de caractères ASCII utilisé habituellement pour les langues européennes méridionales. Également connu sous le nom Latin-3.
ISO-8859-9	Chiffrement de caractères ASCII utilisé habituellement pour la langue turque. Également connu sous le nom Latin-5.
CP850	Une page de code ASCII utilisée pour écrire les langues européennes occidentales.
CP500	Une page de code EBCDIC utilisée pour écrire les langues européennes occidentales.
Shift_JIS	Un encodage de caractères pour la langue japonaise.
MS932	Extension Microsoft de Shift_JIS permettant d'inclure les caractères spéciaux NEC, la sélection NEC des extensions IBM et les extensions IBM.
CP1047	Une page de code EBCDIC avec le jeu de caractères Latin-1 complet.

Paramètres	Description
Délimiteur de champ	<p>Indique le caractère utilisé pour séparer des champs dans un fichier délimité.</p> <p>Par exemple, cet enregistrement utilise une barre verticale () comme délimiteur de champ :</p> <pre>7200 13TH ST MIAMI FL 33144</pre> <p>Les caractères disponibles pour le délimiteur de champ sont les suivants :</p> <ul style="list-style-type: none"> • Virgule • Point-virgule • Conduite • Tabulation • Espace • Point (.)
Qualificateur de texte	<p>Le caractère utilisé pour entourer les valeurs de texte dans un fichier délimité.</p> <p>Par exemple, cet enregistrement utilise des guillemets doubles (") en tant que qualificateur de texte.</p> <pre>"7200 13TH ST" "MIAMI" "FL" "33144"</pre> <p>Ces caractères disponibles à définir en tant que qualificateurs de texte sont :</p> <ul style="list-style-type: none"> • Guillemets simples (') • Guillemets doubles (") <p>Si le fichier utilise un délimiteur de texte, cliquez sur le bouton de sélection pour sélectionner un autre caractère délimiteur.</p>
Séparateur de ligne	<p>Indique le caractère utilisé pour séparer des enregistrements en ligne dans un fichier séquentiel ou délimité.</p> <p>Les paramètres de séparateur d'enregistrements disponibles sont :</p> <p>Unix (U+000A) Un caractère de saut de ligne sépare les enregistrements. Il s'agit du séparateur d'enregistrement standard pour les systèmes Unix.</p> <p>Macintosh (U+000D) Un caractère de retour chariot sépare les enregistrements. Il s'agit du séparateur d'enregistrement standard pour les systèmes Macintosh.</p> <p>Windows (U+000D U+000A) Un retour chariot suivi d'un saut de ligne sépare les enregistrements. Il s'agit du séparateur d'enregistrement standard pour les systèmes Windows.</p>


Paramètres	Description
Première ligne comme en-tête	<p>Indique si le premier enregistrement d'un fichier délimité contient des informations d'en-tête. Un Oui indique qu'il comporte des informations d'en-tête.</p> <p>Par exemple, cet extrait de fichier illustre une ligne d'en-tête dans le premier enregistrement.</p> <pre>"AddressLine1" "City" "StateProvince" "PostalCode" "7200 13TH ST" "MIAMI" "FL" "33144" "One Global View" "Troy" "NY" "12180"</pre>

11. Pour appliquer des paramètres d'un fichier à un autre :

1. Sélectionnez le fichier à partir duquel vous souhaitez copier les paramètres, puis cliquez sur le bouton **Copier les paramètres du fichier**.
2. Dans la fenêtre contextuelle **Appliquer les paramètres fichier** qui s'affiche, sélectionnez les fichiers dans lesquels vous souhaitez appliquer les paramètres, puis cliquez sur **OK**.

Remarque : Les paramètres de fichier sont appliqués au(x) fichier(s) sélectionné(s).

12. Pour afficher un aperçu du fichier sélectionné, reportez-vous à la section **Aperçu** sous **Paramètres de fichiers**.

13. Pour configurer le profil que vous avez créé, cliquez sur  dans le coin supérieur droit de la page.

Ajout de données à partir de modèles

Pour créer un profil à partir de Physical Models ou de Logical Models, procédez comme suit :

1. Dans un navigateur Web, accédez à :

`http://server:port/metadata-insights`

Où : *server* est l'adresse IP ou le nom de serveur de votre serveur Spectrum™ Technology Platform et *port* le port HTTP. Par défaut, le port HTTP est 8080.

2. Connectez-vous à l'aide de vos informations d'identification.
3. Sur la page **Accueil**, cliquez sur **Profilage**.

4. Cliquez sur le bouton **Ajouter profil** .

La page **Ajouter profil** s'affiche avec un assistant affichant les trois phases de création d'un profil : **Sélectionner la source** > **Configurer le profil** > **Terminer**.

5. Dans les options **Sélectionner une source de données**, sélectionnez l'option **Modèle**.


6. Dans la liste déroulante **Sélectionner un modèle**, sélectionnez le Physical Model ou le Logical Model que vous souhaitez utiliser pour créer votre profil. Pour affiner votre recherche par type de modèle (Physical Model ou Logical Model) ou nom de modèle, saisissez-le dans le champ et les modèles associés commencent à s'afficher sous le champ. Par exemple, si vous saisissez « Logical », tous les Logical Models s'affichent.

Remarque : Pour créer un modèle ou pour modifier un modèle existant, cliquez sur le lien **Gérer les modèles**. Il vous permet d'accéder à la page **Modélisation**, où vous pouvez créer et enregistrer un nouveau Physical Model/Logical Model. Ce modèle peut ensuite être sélectionné pour créer un profil à l'aide de la liste déroulante **Sélectionner un modèle**.

Les tables et les colonnes du modèle sélectionné sont affichées sous le champ.

7. Dans la liste des tables, sélectionnez les tables que vous souhaitez inclure dans le profil. Dans chacune des tables, vous pouvez également sélectionner les colonnes dont vous avez besoin pour le nouveau profil. Pour sélectionner toutes les tables et toutes les colonnes, cochez la case de la ligne d'en-tête.

Remarque : Utilisez la zone de texte de recherche à saisie anticipée pour rechercher des tables/colonnes spécifiques.

8. Pour configurer le profil que vous avez créé, cliquez sur  dans le coin supérieur droit de la page.


Ajout de données à partir d'une connexion

Pour créer un nouveau profil à partir de l'une des connexions de base de données configurées dans Spectrum™ Technology Platform, procédez comme suit :

1. Dans un navigateur Web, accédez à :



`http://server:port/metadata-insights`

Où : *server* est l'adresse IP ou le nom de serveur de votre serveur Spectrum™ Technology Platform et *port* le port HTTP. Par défaut, le port HTTP est 8080.

2. Connectez-vous à l'aide de vos informations d'identification.
3. Sur la page **Accueil**, cliquez sur **Profilage**.
4. Cliquez sur le bouton **Ajouter profil** .
La page **Ajouter profil** s'affiche avec un assistant affichant les trois phases de création d'un profil : **Sélectionner la source** > **Configurer le profil** > **Terminer**.
5. Dans les options **Sélectionner une source de données**, sélectionnez l'option **Connexion**.
6. Dans la liste déroulante **Sélectionner une connexion**, sélectionnez la connexion que vous souhaitez utiliser pour créer le profil. Pour affiner votre recherche par type de connexion ou par

son nom, saisissez-le dans le champ et les connexions associées s'affichent sous le champ. Par exemple, si vous saisissez « mssql », toutes les connexions au serveur MSSQL s'affichent.

Remarque : La liste déroulante affiche toutes les connexions qui sont configurées dans Spectrum™ Technology Platform. Pour effectuer une nouvelle connexion, cliquez sur le lien **Gérer connexions...** Vous accédez à la page **Sources de données de Management Console**. Pour connaître les étapes pour configurer une nouvelle source de données, reportez-vous à la section **Data Source Connections** à la page 9.

7. Cliquez sur le bouton **Sélectionner des tables**.
La fenêtre contextuelle **Sélectionner des tables** s'affiche avec toutes les tables de la connexion sélectionnée.
8. Dans la liste des tables de la fenêtre contextuelle **Sélectionner des tables**, sélectionnez les tables que vous souhaitez inclure dans le profil et cliquez sur **OK**.
Les tables sélectionnées sont affichées sous le bouton **Sélectionner des tables** sur la page **Ajouter profil**.
9. Pour sélectionner des colonnes d'une table, cliquez sur  et sélectionnez les colonnes requises pour créer le profil. Pour sélectionner toutes les colonnes, cochez la case de l'en-tête de la table.
Remarque : Utilisez la zone de texte de recherche à saisie anticipée pour rechercher des tables/colonnes spécifiques.
10. Pour configurer le profil que vous avez créé, cliquez sur  dans le coin supérieur droit de la page.

Configuration d'un profil

Pour chaque profil que vous créez, vous devez définir certains paramètres au départ, afin de ne pas avoir à les rechercher chaque fois que vous exécutez l'analyse du profil. Ces paramètres sont les suivants :

- Règles d'analyse de la qualité des données
- Taille d'échantillon pour chaque table du résultat
- Liste des utilisateurs auxquels les résultats de profilage doivent être automatiquement envoyés


En outre, vous pouvez également indiquer le nombre de tables à exécuter simultanément dans le profilage et le nombre d'exécutions qui peuvent être enregistrées pour l'aperçu des données. Pour plus d'informations, voir **Paramètres par défaut pour le profilage**.

Vous pouvez configurer votre profil à l'aide du stage **Configurer le profil** de la page **Ajouter profil**. Pour ce faire, procédez comme suit :



Remarque : Pour connaître les étapes pour accéder à la page **Ajouter profil**, reportez-vous à la section **Ajout de données à partir de fichiers plats** à la page 112.

1. Spécifiez la taille d'échantillon de chaque table à l'aide du champ **Taille de l'échantillon par table**. Par exemple, si vous spécifiez 10 000, l'analyse prend en compte 10 000 lignes de données dans chaque échantillon d'analyse et génère le rapport associé.
2. Dans la liste des **Règles de profilage**, sélectionnez les règles qui sont nécessaires pour le type d'analyse à effectuer sur votre jeu de données. Pour obtenir la liste des règles de profilage et leur description, reportez-vous à la section **Règles de profilage** à la page 119.

Remarque : Vous pouvez filtrer les règles à l'aide de la zone de texte de filtre à saisie anticipée.

3. Dans la section **Notification**, sélectionnez l'option **Quand l'analyse de qualité des données est terminée, envoyer un e-mail à** pour envoyer un courrier électronique contenant le rapport d'analyse à des utilisateurs. Indiquez les adresses électroniques des utilisateurs. Pour ajouter plusieurs adresses électroniques, cliquez sur le bouton **Ajouter** .

Remarque : Pour pouvoir envoyer des courriers automatiques, vous devez configurer un serveur de messagerie à l'aide de **Management Console**. Pour obtenir des informations détaillées sur cette tâche, reportez-vous au *Guide d'administration*.

4. Pour tout type de correction ou de mise à jour, vous pouvez accéder à la phase précédente en cliquant sur  dans le coin supérieur droit de la page.
5. Cliquez sur  dans le coin supérieur droit de la page pour passer à la phase suivante de la création de profil.

Règles de profilage

Les règles de profilage effectuent différents types d'analyse de vos données. Lorsque vous configurez un profil, sélectionnez les règles de profilage qui effectuent les types d'analyse de données qui vous intéressent.

Cette section décrit les règles de profilage prises en charge dans Metadata Insights.


Analyse de caractères

Cette règle identifie les modèles, les scripts et les types de caractère des champs de type chaîne. Lorsque vous activez cette règle, vous obtenez ces informations pour les champs de type chaîne de vos données :

- **Fréquence** : les expressions les plus utilisées dans la colonne de chaîne sélectionnée.
- **Distribution de scripts** : les différents scripts identifiés dans la colonne de chaîne sélectionnée et leur nombre.
- **Longueurs de chaîne** : distribution des longueurs de chaîne dans la colonne de chaîne sélectionnée.
- **Catégories de caractères** : types de caractères de la colonne sélectionnée, tels que les lettres, la ponctuation et les nombres.

- **Motifs de texte** : convertit les données de la colonne de chaîne en motif et affiche le motif, son nombre et son pourcentage d'occurrence. Le motif est déterminé à l'aide de cette règle :
 - Les caractères majuscules latins sont remplacés par « A »
 - Les caractères minuscules latins sont remplacés par « a »
 - Les chiffres sont remplacés par « 9 »
 - Les caractères de contrôle sont remplacés par « c »

Analyse des motifs personnalisés

Cette règle identifie tout motif dans la colonne de chaîne. Vous pouvez configurer autant d'expressions régulières que vous le souhaitez pour effectuer une correspondance avec vos données. Pour configurer cette règle lors de la création d'un profil, cliquez sur l'icône Configurer  et saisissez ces détails :

1. Clé : nom du motif à identifier
2. Valeur : expression régulière du motif

Par exemple, si vous souhaitez identifier les adresses électroniques dans les colonnes de chaînes, saisissez l'expression régulière `^(.+)@(.+) $` dans le champ **Valeur** et e-mail dans le champ **Clé**.

3. Pour ajouter une autre expression, cliquez sur l'icône Ajouter et ajoutez les détails de la paire clé/valeur suivante. Vous pouvez ajouter autant d'expressions que vous le souhaitez pour effectuer une correspondance avec vos données.

Lorsque vous activez cette règle, vous obtenez ces informations :

- **Validité** : valeurs correspondant à au moins l'un des modèles d'expressions régulières de la règle.
- **Répartition des correspondances de modèles** : répartition des enregistrements correspondant aux expressions régulières.

Analyse des dates

Cette règle détecte et valide les dates des colonnes de chaînes. Elle identifie également les modèles de date des colonnes et leur répartition. Cette analyse peut s'avérer utile pour détecter les entrées de date des colonnes erronées, par exemple, dans les données d'e-mail.

Si vous sélectionnez cette règle, la page **Résultats de profilage des données** affiche un onglet **Résumé de date** supplémentaire pour les colonnes de type chaîne qui comportent des dates. Cet onglet affiche ces détails :

- **Validité** : valeurs valides et non valides.
- **Modèles de date** : modèles de date détectés dans les colonnes sélectionnées, leur nombre total et le pourcentage de ce modèle dans le jeu de données.

Analyse des courriers électroniques

Cette règle détecte et valide les adresses électroniques et détermine la répartition des domaines électroniques dans la colonne de données sélectionnée.

Si vous sélectionnez cette règle, la page **Résultats de profilage des données** affiche un onglet **Résumé de l'email** supplémentaire affichant ces détails :

- **Validité** : valeurs valides et non valides.
- **Distribution de domaines** : les dix premiers domaines électroniques de la colonne de données sélectionnée.

Analyse des numéros de téléphone

Sélectionnez cette règle pour détecter et valider les numéros de téléphone et identifier les numéros de téléphone comme les numéros de ligne fixe, les numéros de téléphone portable ou tout autre type de numéro. Cette règle fournit également la répartition des numéros de téléphone par pays et par région. Vous devez configurer cette règle pour définir le pays par défaut à utiliser quand un numéro de téléphone ne comporte pas de code de pays.

Si vous sélectionnez cette règle, la page **Résultats de profilage des données** affiche un onglet **Résumé du numéro de téléphone** supplémentaire affichant ces détails :

- **Validité** : numéros de téléphone valides et non valides.
- **Types de numéro de téléphone** : types de numéro de téléphone tels que téléphone mobile, ligne terrestre, ligne fixe, VOIP, radiomessagerie, messagerie vocale ou numéro sans frais.
- **Numéros de téléphone par pays** : répartition des numéros de téléphone détectés par pays.
- **Numéros de téléphone par région** : répartition des numéros de téléphone détectés par région.


Analyse sémantique

Sélectionnez cette règle pour détecter les types de sémantique, tels que le prénom, la ville, le pays, les codes de pays ISO 2 et 3, le nom de famille et l'état. Cette règle peut vous aider à trouver des valeurs dans des colonnes incorrectes, telles que les noms de ville d'une colonne Pays.

Si vous sélectionnez cette règle, la page **Résultats de profilage des données** affiche un onglet **Type de sémantique** supplémentaire affichant les types de sémantique détectés et leur fréquence.

Analyse des adresses américaines

Cette règle détermine la qualité de vos données d'adresse à l'aide de la base de données U.S. du module Universal Addressing. Pour lancer cette règle, vous devez :

- Installer la base de données U.S. du module Universal Addressing et la définir comme ressource dans Management Console. Pour plus d'informations sur l'ajout de cette ressource de base de données, reportez-vous au *Guide d'administration*.
- Configurez la règle Analyse des adresses américaines en cliquant sur le bouton Configurer  et en saisissant ces informations :
 1. **Base de données US Address Coder** : sélectionnez la ressource de base de données du module Universal Addressing configurée dans Management Console.
 2. **Champ AddressLine1 à Champ AddressLine5** : mappez ces champs vers les colonnes de la table que vous analysez. Vous n'avez pas nécessairement besoin de saisir des noms de

colonne dans tous les champs. Cependant, plus vous serez spécifique, meilleur sera le score de correspondance.


3. Mappez les colonnes de votre table vers les champs **City**, **Country**, **USUrbanName**, **FirmName field**, **PostalCode** et **StateProvince**.

Si vous sélectionnez cette règle, la page **Résultats de profilage des données** affiche l'onglet **Résumé de l'adresse**.

- La légende sous le graphique affiche le score de correspondance des données ainsi que le codage couleur.
- Pointez n'importe où dans la zone du graphique pour afficher le score de correspondance. Les scores se présentent sous forme de plages (0, 1-25, 26-50, 51-80, 81-99 et 100), avec zéro représentant l'absence de correspondance des données par rapport à la base de données. Le graphique affiche également le pourcentage d'enregistrements correspondants détectés (codés par couleur).
- Cliquez sur la zone du graphique pour afficher les données correspondant ou non à la base de données.

Analyse des adresses internationales

Cette règle détermine la qualité de vos données d'adresse à l'aide de la base de données du module Global Address Validation. Pour lancer cette règle, vous devez :

- Installez la base de données Global Address Validation et définissez-la comme ressource dans Management Console. Pour plus d'informations, reportez-vous au *Guide d'administration*.
- Configurez la règle Analyse des adresses internationales en cliquant sur le bouton Configurer  et en saisissant ces informations :
 1. **Base de données Addressing Engine** : sélectionnez la ressource de base de données Global Address Validation configurée dans **Management Console**.
 2. **Champ AddressLine1** et champ **Country** : mappez ces champs vers les colonnes de la table que vous analysez.
 3. Mappez les colonnes de votre table vers les champs **LastLine**, **City**, **CitySubdivision**, **PostalCode**, **State**, **StateSubdivision** et **FirmName**. Vous n'avez pas nécessairement besoin de saisir des noms de colonne dans tous les champs. Cependant, plus vous serez spécifique, meilleur sera le score de correspondance.

Si vous sélectionnez cette règle, la page **Résultats de profilage des données** affiche l'onglet **Résumé de l'adresse**, qui affiche :

- **Distribution en confiance des adresses internationales** : score de correspondance des données. Le score est codé par couleur. Pointez n'importe où dans le graphique en forme de donut pour afficher les plages de scores (0, 1-25, 26-50, 51-80, 81-99 et 100), avec zéro représentant l'absence de correspondance des données dans la base de données. Cliquez sur la zone du graphique pour obtenir un aperçu des données en correspondance ou non.
- **Précision des adresses internationales** : répartition des niveaux de validation des adresses, par exemple, état, résidence, code postal, ville, sous-division de ville et rue.


Définition d'un profil

La troisième et dernière phase de la création du profil consiste à définir le profil, à savoir, lui attribuer un nom et une description. Pour ce faire, à la phase Terminer de la page **Ajouter profil**, procédez comme suit :

Remarque : Pour connaître les étapes pour accéder à la page **Ajouter profil**, reportez-vous à la section **Ajout de données à partir de fichiers plats** à la page 112.

1. Saisissez un nom pour le profil dans le champ **Nom de profil**.
2. Décrivez le profil dans le champ **Description**.

Remarque : La description peut inclure l'objectif de la création du profil.

3. Pour tout type de correction ou de mise à jour, vous pouvez accéder à la phase précédente en cliquant sur  dans le coin supérieur droit de la page.
4. Pour enregistrer le profil et l'exécuter ultérieurement, cliquez sur **Enregistrer**.
Le profil est créé et ajouté à la page **Profilage**.
5. Pour enregistrer et exécuter le profil, cliquez sur **Enregistrer et exécuter**.
Le profil est créé et ajouté à la page **Profilage** et le processus d'analyse commence dessus.
Vous pouvez afficher des détails, comme l'état d'exécution du profil, sur la page **Profilage**.

Spécification des paramètres de profilage par défaut

Pour les profils que vous avez configurés, vous pouvez modifier ces deux paramètres :

- Le nombre de tables pouvant être exécutées simultanément lors de l'exécution du profil
- Le nombre d'exécutions du profil pendant lequel les données des fichiers doivent être disponibles pour l'aperçu

Ces règles vous permettent d'optimiser les performances en fonction de vos besoins et de la taille des fichiers. Pour spécifier ces paramètres, procédez comme suit :

1. Dans un navigateur Web, accédez à :

`http://server:port/metadata-insights`

Où : *server* est l'adresse IP ou le nom de serveur de votre serveur Spectrum™ Technology Platform et *port* le port HTTP. Par défaut, le port HTTP est 8080.

2. Connectez-vous à l'aide de vos informations d'identification.
3. Sur la page **Accueil**, cliquez sur **Profilage**.

La page **Profilage** apparaît.

4. Cliquez sur le lien **Valeurs par défaut**.
La page **Options de profilage par défaut** s'affiche.
5. Dans la case **Nombre de tables du profilage simultan **, sp cifiez le nombre de tables qui doivent  tre ex cut es simultan ment pour l'analyse.

Remarque : Ce param tre s'applique   tous les profils que vous d finissez pour qu'ils soient ex cut s. La plage de tables valide est de 1   20, avec une valeur par d faut  gale   5.

Exemple : **Nombre de tables du profilage simultan ** est d fini sur 5 et deux profils, chacun comportant six tables, doivent  tre ex cut s :

Metadata Insights commence par ex cuter cinq tables du profil programm  pour  tre ex cut  et conserve le deuxi me profil dans la file d'attente (l' tat est affich  dans le champ ** tat**, de l'onglet **Profils** de la page **Profilage**). Les autres tables sont s lectionn es l'une apr s l'autre   mesure que l'ex cution du profil termine les tables s lectionn es.

6. Dans la case **Conserver les donn es sources pour [N] ex cutions**, sp cifiez le nombre d'ex cutions pour lesquelles les tables s lectionn es dans l'analyse du profil sont enregistr es dans l'aper u des donn es. La valeur par d faut est 2.

Exemple : **Conserver les donn es sources pour [N] ex cutions** est d fini sur 5.

Metadata Insights n'affiche pas les donn es des tables dans l'aper u apr s avoir termin  cinq ex cutions pour celles-ci. Si vous cliquez sur les graphiques, les barres ou les graphiques en forme de donut correspondant   ces tables sur la page **R sultats de profilage des donn es**, vous obtenez le message : *L'aper u n'est pas disponible*. Cependant, les graphiques, les barres et les graphiques en forme de donut sont affich s, quelles que soient les ex cutions sp cifi es dans ce champ.

Analyse d'un profil

Une fois un profil cr e et r pertori  dans l'onglet **Profils**, vous pouvez l'envoyer pour qu'il soit analys . Le profil est analys  en fonction des r gles s lectionn es et un rapport s'affiche sur la page **R sultats de profilage des donn es**. Le rapport peut  galement  tre envoy  directement aux utilisateurs cibl s si vous le sp cifiez sur la page **Ajouter profil** et vous pouvez  galement le t l charger au format PDF.

L'analyse du profil peut  tre ex cut e comme une activit  ponctuelle ou un  v nement r current en programmant la r currence sur la page **Ajouter un calendrier**. Vous pouvez  galement d marrer l'analyse directement   partir de l'onglet **Profils**. Dans tous ces cas, l'analyse est ex cut e et le statut affich  dans les onglets **Profils** et **Historique**.

Quand un profil commence à être analysé, un ID unique lui est affecté, qui s'affiche dans l'onglet **Historique**. Pour plus d'informations, reportez-vous à la section [Affichage de l'historique des analyses](#) à la page 128.


Remarque : Si le même profil est envoyé pour être analysé plusieurs fois, à chaque analyse est affecté un ID différent.

Pour exécuter l'analyse de profil à partir de la liste des profils affichés dans l'onglet **Profils**, procédez comme suit :

1. Dans un navigateur Web, accédez à :

`http://server:port/metadata-insights`

Où : *server* est l'adresse IP ou le nom de serveur de votre serveur Spectrum™ Technology Platform et *port* le port HTTP. Par défaut, le port HTTP est 8080.

2. Connectez-vous à l'aide de vos informations d'identification.
3. Sur la page **Accueil**, cliquez sur **Profilage**.
4. Dans la liste des profils affichés, cochez la case correspondant à celui que vous souhaitez analyser, puis cliquez sur l'icône **Exécuter profil** .

Remarque : Vous pouvez rechercher le profil requis à l'aide de la zone de texte à saisie anticipée **Filtre**. Elle vous permet d'effectuer une recherche en fonction du Nom, de l'Utilisateur ou de la Source de données du profil.


Le jeu de données sélectionné du profil commence à être analysé, avec la colonne **Statut** affichant l'icône de progression. L'icône se transforme en une coche lorsque le profilage est terminé. Les autres détails du profil affichés sont les suivants :

- **Nom** : nom du profil. Un clic sur le nom du profil vous permet d'accéder à la page **Résultats de profilage des données**. Pour plus d'informations, reportez-vous à la section [Détails des tables et des colonnes dans les résultats de profilage](#) à la page 130.
- **Source de données** : source de données du profil en cours d'analyse. Il peut s'agir d'un Physical Model ou d'un Logical Model ou encore d'une Data Source Connection directe.
- **Utilisateur** : l'utilisateur qui a créé le profil
- **Heure de début de la dernière exécution** : date (au format mm/jj/aa) et heure (au format hh:mm et AM/PM) de démarrage de l'analyse de profil la plus récente
- **Heure de fin de la dernière exécution** : date (au format mm/jj/aa) et heure (au format hh:mm et AM/PM) de fin de l'analyse de profil la plus récente
- **Durée**: durée (au format mm:ss) de l'analyse
- **Statut** : si l'analyse est en cours, terminée ou si elle a échoué
- **Tables** analysées
- **Enregistrements** analysés
- **Plus** : cliquez sur l'icône de PDF pour télécharger le rapport d'analyse. Dans le cas où le profilage rencontre des obstacles, vous pouvez afficher l'erreur dans ce champ.

Programmation de l'analyse

Pour planifier une analyse de profil, utilisez la page **Ajouter un calendrier**. Cette page vous offre différentes combinaisons de plages de modèles, d'heures et de dates pour programmer le profilage en fonction de vos besoins. Vous pouvez également effectuer cette activité de manière ponctuelle.

Pour planifier l'analyse d'un profil, procédez comme suit :

1. Sur la page **Profilage**, cliquez sur l'onglet **Calendriers**, puis sur l'icône **Ajouter des calendriers** .
2. La page **Ajouter un calendrier** s'affiche.
3. Dans le champ **Nom**, saisissez le nom du calendrier.
4. Dans la liste déroulante **Profil**, sélectionnez le profil sur lequel vous souhaitez programmer une analyse.

Remarque : Tous les profils définis sont affichés.

4. Dans la section **Calendrier**, sélectionnez l'un des déclencheurs (dans la liste déroulante **Déclencheur**), selon les besoins.
 - **Date/heure** : pour planifier une analyse unique du profil. Pour cette analyse, sélectionnez la **Date de début** (au format jj/mm/aa) et l'**Heure de début** (au format hh/mm et AM/PM) dans les calendriers respectifs.
 - **Date/heure récurrentes** : pour planifier une analyse récurrente du profil. La récurrence peut être définie sous forme de **Modèle** (de jours, semaine ou mois), de **Temps** (par exemple, toutes les trois heures), de **Plage** (à un moment indiqué entre une plage de dates) ou d'une combinaison des trois. Pour plus d'informations, reportez-vous à la section **Programmation de l'analyse récurrente** à la page 127.
5. Pour activer ce calendrier, assurez-vous que le curseur **Calendrier** est maintenu sur **Activé**.
6. Cliquez sur **Enregistrer**.

Le calendrier de profilage est enregistré et affiché dans l'onglet **Calendriers** de la page **Profilage**. Cette page répertorie tous les profils programmés pour être analysés et affiche ces détails pour chacun :

- **Profil** : nom du profil
- **Nom du calendrier** : nom du calendrier
- **Utilisateur** : utilisateur qui a créé le calendrier
- **Déclencheur** : il spécifie **CALENDRIER** pour indiquer que le profil est programmé pour être analysé
- **Exécutez suivante** : indique la date (au format aaaa/mm/jj), l'heure (au format hh/mm/ss) et la période (AM/PM) d'exécution suivante du calendrier.
- **Dernière exécution** : indique la date (au format aaaa/mm/jj), l'heure (au format hh/mm/ss) et la période (AM/PM) de la dernière exécution du calendrier.

Remarque : Lorsque l'analyse d'un profil commence, le profil en cours s'affiche également dans l'onglet **Profils** de la page **Profilage** s'il s'agit d'une programmation récurrente. Si la programmation n'est pas récurrente, le profil en cours est déplacé de l'onglet **Planification** de l'onglet **Profils** lors du démarrage du traitement.

Programmation de l'analyse récurrente

Vous pouvez définir la récurrence de l'analyse de profil à l'aide de la combinaison requise de **Modèle**, **Temps** et **Plage** dans la section **Calendrier** de la page **Ajouter un calendrier**.

Motif

- **Chaque jour de la semaine** : pour exécuter l'analyse chaque jour de la semaine
- **Tous les** : à une périodicité spécifiée. Par exemple, pour exécuter une analyse tous les deux jours, sélectionnez cette option, puis saisissez 2 dans la zone de sélection.
- **Jour <> de chaque <> mois** : pour exécuter l'analyse un jour et un mois donnés. Par exemple, pour exécuter l'analyse le troisième jour tous les deux mois, saisissez 3 dans la zone de sélection **Jour** et 2 dans la zone de sélection **Mois**.
- **Tous les <> Mois <un jour spécifié> <de la semaine>** : par exemple, pour exécuter l'analyse le deuxième mardi tous les trois mois, saisissez 3 dans la zone de sélection **Mois**, et, dans les listes déroulantes, sélectionnez les options **Le deuxième** et **Mardi**, respectivement.
- **Jours de la semaine** : sélectionnez les jours de la semaine durant lesquels vous souhaitez exécuter l'analyse.

Remarque : Vous pouvez sélectionner plusieurs jours.

Durée

Pour exécuter l'analyse à un intervalle de temps spécifique, sélectionnez cette option, et, dans les zones de sélection, spécifiez le nombre d'heures ou minutes après lesquelles vous souhaitez exécuter l'analyse.

Plage

Pour spécifier la plage d'analyse, sélectionnez la **Date de début** et la **Date de fin** (au format jj/mm/aa) dans les calendriers respectifs et sélectionnez l'heure (au format hh:mm) et la période (AM/PM) dans l'option **Heure de début**.

Remarque : Cochez la case **Aucune date de fin** si vous ne souhaitez pas définir une date de fin pour l'analyse récurrente.

Affichage de l'historique des analyses

L'onglet **Historique** de la page **Profilage** affiche une liste de toutes les analyses exécutées sur différents profils avec des détails, tels que l'heure de début et l'heure de fin de l'analyse, si elle a été correctement effectuée, l'ID unique affecté à l'analyse par le système, le nombre de tables et d'enregistrements analysés dans chaque cas et la source de données des profils analysés.

Remarque : Il est recommandé de continuer à supprimer les exécutions de profil obsolètes de l'historique.

Pour afficher l'historique de toutes les analyses exécutées par l'application, procédez comme suit :

1. Dans un navigateur Web, accédez à :

`http://server:port/metadata-insights`


Où : *server* est l'adresse IP ou le nom de serveur de votre serveur Spectrum™ Technology Platform et *port* le port HTTP. Par défaut, le port HTTP est 8080.

2. Connectez-vous à l'aide de vos informations d'identification.
3. Sur la page **Accueil**, cliquez sur **Profilage**.
4. Cliquez sur l'onglet **Historique**.

Une liste de toutes les analyses exécutées sur les différents profils s'affiche, avec ces détails :



- **ID** : identification unique attribuée par l'application à cette analyse de profil.

Remarque : Dans le cas où le même profil est analysé plusieurs fois, chaque exécution dispose d'un ID unique.

- **Nom** : nom du profil. Un clic sur le nom du profil vous permet d'accéder à la page **Résultats de profilage des données**. Pour plus d'informations, reportez-vous à la section **Détails des tables et des colonnes dans les résultats de profilage** à la page 130.
 - **Source de données** : source de données du profil en cours d'analyse. Il peut s'agir d'un Physical Model ou d'un Logical Model ou encore d'une Data Source Connection directe.
 - **Utilisateur** : l'utilisateur qui a créé le profil
 - **Heure de début de la dernière exécution** : date (au format mm/jj/aa) et heure (au format hh:mm et AM/PM) de démarrage de l'analyse de profil la plus récente
 - **Heure de fin de la dernière exécution** : date (au format mm/jj/aa) et heure (au format hh:mm et AM/PM) de fin de l'analyse de profil la plus récente
 - **Durée**: durée (au format mm:ss) de l'analyse
 - **Statut** : si l'analyse est en cours, terminée ou si elle a échoué
5. Pour afficher le rapport de profilage, sélectionnez le profil requis, puis cliquez sur l'icône **Afficher l'historique** .

Remarque : Utilisez la zone de texte à saisie anticipée **Filtre** pour rechercher le profil plus rapidement. Elle vous permet d'effectuer une recherche en fonction des détails des champs Nom, Source de données et Utilisateur.

La page **Résultats de profilage des données** s'affiche. Pour plus d'informations, reportez-vous à la section **Détails des tables et des colonnes dans les résultats de profilage** à la page 130.

6. Pour actualiser la page, cliquez sur l'icône **Actualiser l'historique** .
7. Pour supprimer une analyse, sélectionnez-la, puis cliquez sur l'icône **Supprimer l'historique** .

Affichage de rapports d'analyse du profil

La page **Résultats de profilage des données** affiche une analyse complète d'un profil, notamment l'exhaustivité et l'unicité des données contenues. Le rapport est généré au niveau du profil, ainsi qu'au niveau des tables et des colonnes. Au niveau des colonnes, le résultat affiché varie suivant les règles de profilage que vous avez sélectionnées. Par exemple, si vous sélectionnez la règle **Analyse sémantique** sur la page **Ajouter profil**, les **Résultats de profilage des données** affichent les types de sémantique détectés dans les colonnes de chaînes sélectionnées.

Pour accéder aux **Résultats de profilage des données**, procédez comme suit :

1. Dans un navigateur Web, accédez à :

`http://server:port/metadata-insights`

Où : *server* est l'adresse IP ou le nom de serveur de votre serveur Spectrum™ Technology Platform et *port* le port HTTP. Par défaut, le port HTTP est 8080.

2. Connectez-vous à l'aide de vos informations d'identification.
3. Sur la page **Accueil**, cliquez sur **Profilage**.
4. Effectuez l'une des actions suivantes :
 - Dans l'onglet **Profils**, cliquez sur le **Nom** du profil dont vous souhaitez afficher le rapport d'analyse.

Remarque : Le **Nom** du profil se présente au format hypertexte uniquement si son analyse a été exécutée.
 - Dans l'onglet **Profils**, cliquez sur l'icône de document de la colonne **Plus** correspondant au profil requis.
 - Dans l'onglet **Historique**, cliquez sur l'ID du profil dont vous souhaitez afficher le rapport d'analyse.

La page **Résultats de profilage des données** s'affiche avec les résultats de l'analyse.

Détails des tables et des colonnes dans les résultats de profilage

La page **Résultats de profilage des données** peut être divisée en trois parties.

- Un en-tête affichant les détails du profil
- Le volet gauche, qui affiche le nom du profil et toutes les tables et colonnes incluses
- Le volet de rapport, qui fournit une vue illustrée des données analysées

Le rapport affiché dans le volet droit est basé sur les éléments que vous sélectionnez dans le volet gauche (profil, table ou colonne). Des détails, tels que les valeurs Déviation standard (**Dév. std**), Maximum (**Maxi**), Minimum (**Mini**) et **Moyenne** sont affichés au niveau des colonnes pour les données numériques, tandis que la fréquence de longueur et la fréquence de motif de texte sont affichées pour les colonnes de chaînes. En outre, suivant les règles sélectionnées pour le profilage, d'autres onglets sont affichés dans le rapport, tels que **Type de sémantique** pour les colonnes comportant des valeurs de type de sémantique si la règle **Analyse sémantique** est sélectionnée. Pour plus d'informations, reportez-vous à la section **Règles de profilage** à la page 119.

Détails du profil

L'en-tête du rapport affiche ces détails du profil analysé.

- **Nom de profil** : nom du profil
- **Description** : brève description du profil
- **Démarré** : date (aaaa/mm/jj), heure (hh:mm:ss) et période (AM ou PM) d'initialisation de l'analyse du profil
- **Source de données** : source de données du profil en cours d'analyse. Il peut s'agir d'un Physical Model ou d'un Logical Model ou encore d'une Data Source Connection directe.
- **Utilisateur** : l'utilisateur qui a créé le profil
- **Tables** : nombre de tables du modèle sélectionné
- **Type de source de données** : indique si la source de données du profil est un modèle ou une base de données.
- **Enregistrements** : nombre total d'enregistrements analysés

Détails du modèle et des tables

Le volet de navigation gauche de la page **Résultats de profilage des données** affiche le profil analysé et les tables et les colonnes qui y sont incluses. Le rapport illustré affiché dans le volet droit correspond à la sélection effectuée dans le volet gauche. Par exemple, si vous cliquez sur le nom de profil, le rapport affiche :

- Le **Résumé** du profil affichant le pourcentage de lignes complètes et incomplètes des données du profil.
- Un aperçu des tables des données du profil dans l'onglet **Résumé des tables**.

Vous pouvez utiliser la zone de texte à saisie anticipée **Rechercher** au-dessus de la grille **Résumé** pour rechercher la table ou la colonne pertinente.

Remarque : Pour le type de données `date`, les résultats affichent le fuseau horaire du poste client.

Les résultats de profilage des données

Suivant les éléments que vous sélectionnez dans le volet gauche de cette page, ces rapports sont affichés dans le volet droit.

- **Rapport d'analyse du profil** : cliquez sur le nom de source de données dans le volet gauche pour afficher ces détails :
 - **Complétude** : pourcentage d'exhaustivité des lignes des données du profil
 - **Résumé des tables** : affiche ces détails pour chaque table du profil
 - **Nom de la table** : noms de toutes les tables du profil
 - **Nombre de colonnes** : nombre de colonnes de la table
 - **Nombre d'enregistrements** : nombre d'enregistrement de la table
 - **Complétude (%)** : exhaustivité des enregistrements de la table
- **Rapport de la table** : cliquez sur l'un des noms de table du volet gauche pour afficher ces détails :
 - **Complétude** : pourcentage d'exhaustivité des lignes de la table
 - **Résumé des colonnes** : affiche ces détails pour chaque colonne de la table
 - **Nom de colonne** : noms de toutes les colonnes de la table
 - **Type de colonne** : type de données de la colonne
 - **Complétude (%)** : exhaustivité des enregistrements de la colonne
 - **Unicité (%)** : caractère unique des données contenues dans la colonne
 - **Type détecté** : affiche les types de sémantique, tels que l'adresse électronique, le téléphone, la ville, le prénom et le nom de famille, détectés dans la chaîne de cette colonne
 - **Fréquence de comptage nulle** : affiche le nombre de valeurs null de chaque ligne de la table
- **Rapport de la colonne** : cliquez sur l'une des colonnes du volet gauche pour afficher ces détails :
 - **Complétude** : pourcentage d'exhaustivité des données de la colonne
 - **Unicité** : caractère unique des données de la colonne (%)
 - **Longueur mini.** et **Longueur maxi.** de caractères au cas où les données de la colonne constituent une chaîne
 - Déviation Standard (**Dév. std**), valeurs maximale (**Maxi.**), minimale (**Mini.**) et **Moyenne** si la colonne comporte des données numériques.
 - **Fréquence** des données de la colonne
 - **Fréquence de longueur** au cas où les données de la colonne sont de type chaîne
 - **Fréquence de modèle de texte** au cas où les données de la colonne sont de type chaîne

4 - Généalogie et analyse de l'impact

In this section

Affichage de la généalogie et de l'analyse de l'impact	133
Icônes dans Généalogie et Analyse de l'impact	135
Utilise	140

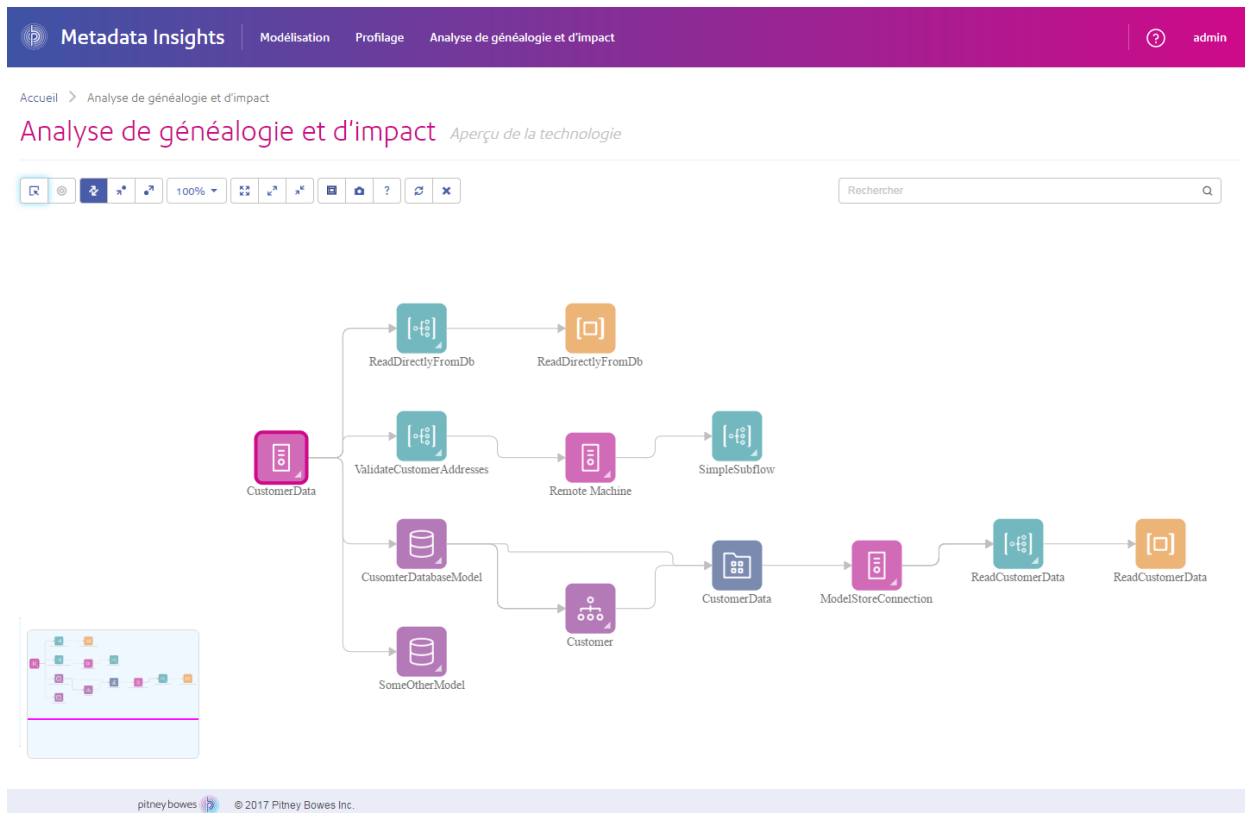
Affichage de la généalogie et de l'analyse de l'impact

La vue Analyse de généalogie et d'impact affiche la manière dont les données fluent des sources de données vers les destinations de données et à travers les flux Spectrum™ Technology Platform. La généalogie et l'analyse de l'impact sont des concepts similaires qui décrivent différentes manières de tracer le flux de données.

Généalogie indique la provenance des données. Vous pouvez l'utiliser pour tracer le chemin d'accès aux données depuis leur source, en affichant tous les systèmes qui traitent et stockent les données en cours de route, tels que les flux Spectrum™ Technology Platform, les bases de données et les fichiers.

Analyse d'impact indique la destination des données et les systèmes qui dépendent des données d'une ressource de données sélectionnée. Vous pouvez l'utiliser pour afficher les flux, les bases de données et les fichiers qui utilisent une ressource de données directement ou indirectement. L'analyse de l'impact s'avère utile si vous souhaitez comprendre comment une modification d'une base de données, d'un fichier ou d'un flux affecte les processus et les systèmes qui utilisent les données.


Metadata Insights peut afficher la généalogie et l'analyse de l'impact sur un seul diagramme qui montre le flux de données complet, de la source à la destination. Vous pouvez également choisir d'afficher uniquement la généalogie ou l'impact. En affichant ensemble la généalogie et l'analyse de l'impact des données, vous pouvez repérer les problèmes de vos processus de données et planifier des mises à niveau et des modifications à apporter à ces processus.





1. Dans un navigateur Web, accédez à :

`http://server:port/metadata-insights`

Où : *server* est l'adresse IP ou le nom de serveur de votre serveur Spectrum™ Technology Platform et *port* le port HTTP. Par défaut, le port HTTP est 8080.

2. Connectez-vous.
3. Cliquez sur **Analyse de généalogie et d'impact**.
4. Cliquez sur le bouton Sélectionner une entité .
5. Sélectionnez le type d'entité à analyser. Par exemple, si vous souhaitez afficher l'analyse de généalogie ou d'impact d'un flux de données, sélectionnez **Flux de données**.
6. Sélectionnez l'entité spécifique dont vous souhaitez analyser les données.
7. Cliquez sur **OK**.

Le schéma qui en résulte affiche les flux de données entrant et sortant de l'entité que vous avez sélectionnée. La généalogie de l'entité figure à gauche de l'entité. L'impact de l'entité figure à droite. Vous pouvez modifier le diagramme pour qu'il affiche uniquement la généalogie de l'entité  ou uniquement l'impact de l'entité .

Certaines entités peuvent être développées, ce qui vous permet d'afficher le flux de données au sein de l'entité, comme les tables spécifiques via lesquelles les données fluent dans une base de

données, ou les stages d'un flux Spectrum™ Technology Platform. Pour développer une entité, cliquez sur l'onglet dans le coin inférieur droit de l'icône de l'entité :



Une fois l'entité développée, vous pouvez afficher les détails du flux de données de l'entité. Par exemple, nous avons développé ici l'entité de base de données nommée Prod SQL Server. Vous pouvez voir que les données de la table Customer Table proviennent du flux Spectrum™ Technology Platform nommé Add to Customer DB. En outre, vous pouvez voir que les données de la table Customer Table sont utilisées par deux flux : Query Customer Job et Geocode Job. Notez également que le flux Geocode Job écrit des données en retour dans Customer Table. Voici donc un exemple dans lequel Customer Table est à la fois une entrée et une sortie de Geocode Job. Cet exemple peut représenter une situation dans laquelle Geocode Job lit les adresses des clients de Customer Table, détermine les coordonnées de latitude et de longitude des adresses, puis ajoute les champs de latitude et de longitude aux enregistrements des clients dans Customer Table.




Icônes dans Généalogie et Analyse de l'impact


Le canevas de Généalogie et Analyse de l'impact de Metadata Insights contient une grande variété d'icônes qui représentent différentes parties de Spectrum™ Technology Platform ainsi que les ressources de données connectées à Spectrum™ Technology Platform.

Flux de données et stages


Icône	Entité	Généalogie	Analyse de l'impact
	<p>Type de flux de données</p> <p>Développez cette icône pour afficher les versions exposées et les dernières versions enregistrées du flux de données.</p>	<p>La généalogie d'un type de flux de données affiche les sources de données fournies à toute version du flux de données (qu'il s'agisse de celle exposée ou de la dernière enregistrée).</p>	<p>L'analyse de l'impact d'un type de flux de données affiche les destinations auxquelles toute version du flux de données écrit des données. Si n'importe quelle version du flux de données est un job exposé, l'analyse de l'impact affiche également une icône de type d'activité qui indique le fait que le flux de données peut être utilisé comme une activité dans un flux de processus.</p>
	<p>Versión du flux de données (exposée ou la dernière enregistrée)</p> <p>Développez cette icône pour afficher les stages du flux de données.</p>	<p>La généalogie d'une version de flux de données affiche les sources de données fournies à la version. Elle affiche également les sous-flux contenus dans le flux de données.</p>	<p>L'analyse de l'impact d'une version de flux de données affiche les destinations auxquelles la version sélectionnée écrit des données. Si le flux de données est un job exposé, l'analyse de l'impact affiche également une icône de type d'activité qui indique le fait que le flux de données peut être utilisé comme une activité dans un flux de processus.</p>
	<p>Type de stage</p>	<p>Les types de stage comportent une généalogie si le stage représente un sous-flux. Dans ce cas, le sous-flux s'affiche comme la généalogie du type de stage. Pour les autres types de stage, il n'existe aucune généalogie.</p>	<p>L'analyse de l'impact d'un type de stage affiche tous les flux qui utilisent le stage. Par exemple, si vous recherchez l'analyse de l'impact du type de stage Read from File, celle-ci affiche tous les flux contenant un stage Read From File.</p>
	<p>Instance de stage</p>	<p>La généalogie d'une instance de stage affiche les données utilisées par le stage ainsi que le type de stage.</p>	<p>L'analyse de l'impact d'une instance de stage indique où vont les données une fois traitées par le stage. Il peut s'agir d'un autre stage d'un flux de données, d'une base de données, d'un fichier ou de tout autre emplacement où les données sont envoyées.</p>

Icône	Entité	Généalogie	Analyse de l'impact
	Options de service par défaut Cette entité représente les options par défaut du service sélectionné. Les options par défaut sont celles configurées dans Management Console.	La généalogie des options de service par défaut affiche les entités sur lesquelles le service s'appuie par défaut. Par exemple, si un service dispose d'une option par défaut qui spécifie une base de données Spectrum comme base de données postales ou base de données de géocodage, cette base de données Spectrum est indiquée dans la généalogie des options de service par défaut.	L'analyse de l'impact des options de service par défaut affiche les stages qui sont configurés pour utiliser les options de service par défaut au lieu de les remplacer dans le stage.



Sources de données et destinations

Icône	Entité	Généalogie	Analyse de l'impact
	Connexion	La généalogie d'une connexion affiche les entités qui écrivent des données vers la connexion.	L'impact d'une connexion affiche les entités qui lisent des données depuis la connexion.



Bases de données Spectrum



Icône	Entité	Généalogie	Analyse de l'impact
	Base de données Spectrum	Les bases de données Spectrum n'ont aucune généalogie.	L'analyse de l'impact des bases de données Spectrum affiche les stages qui utilisent la base de données. Elle indique également lorsque la base de données est configurée comme une option de service par défaut.

Rapports





Icône	Entité	Généalogie	Analyse de l'impact
	Type de rapport	Les types de rapport n'ont aucune généalogie.	L'analyse de l'impact d'un type de rapport affiche les flux dans lesquels le type de rapport est utilisé. Par exemple, si vous recherchez l'analyse de l'impact du type de rapport ValidateAddressSummary, celle-ci affiche tous les flux qui génèrent un rapport ValidateAddressSummary.
	Instance de rapport	La généalogie d'une instance de rapport affiche le type de rapport.	Les instances de rapport n'ont aucun impact.

Flux de processus et activités

Icône	Entité	Généalogie	Analyse de l'impact
	Type de flux de processus Développez cette icône pour afficher les versions exposées et les dernières versions enregistrées du flux de processus.	La généalogie d'un type de flux de processus affiche les activités, telles que des flux de données, qui sont appelées par n'importe quelle version du flux de processus, et les sources de données fournies à ces activités. La sortie des activités s'affiche également dans la généalogie du flux de processus, mais pas dans l'impact.	Les types de flux de processus n'ont aucun impact dans la mesure où toutes les données produites suite à l'exécution d'un flux de processus sont affichées dans la généalogie, où les activités appelées apparaissent avec leur sortie.
	Version de flux de processus Développez cette icône pour afficher les activités du flux de processus.	La généalogie d'une version de flux de processus affiche les activités, telles que des flux de données, qui sont appelées par la version sélectionnée du flux de processus, et les sources de données fournies à cette version.	Les versions de flux de processus n'ont aucun impact dans la mesure où toutes les données produites suite à l'exécution d'un flux de processus sont affichées dans la généalogie, où les activités appelées apparaissent avec leur sortie.

Icône	Entité	Généalogie	Analyse de l'impact
	Type d'activité Un type d'activité représente un flux qui peut être utilisé à l'intérieur d'un flux de processus.	La généalogie d'un type d'activité affiche le flux de données représenté par l'activité et les sources de données du flux de données.	L'impact d'un type d'activité affiche les flux de processus qui contiennent l'activité. Si un type d'activité n'a aucun impact, il n'est utilisé dans aucun flux de processus.
	Instance d'activité	La généalogie d'une instance d'activité est le type d'activité et le flux de données représenté par le type d'activité. Si l'activité est précédée d'une autre activité dans un flux de processus, cette activité est également affichée dans la généalogie de l'activité.	L'impact d'une activité est l'activité qui la suit dans un flux de processus.

Modeling

Icône	Entité	Généalogie	Analyse de l'impact
	Model store	La généalogie d'un Model Store affiche les Physical Models et les Logical Models qui composent le Model Store.	L'impact d'un Model Store affiche les connexions vers le Model Store et les flux qui utilisent le Model Store.
	Physical Model	La généalogie d'un Physical Model affiche les connexions qui fournissent des données au Physical Model, telles que les connexions aux bases de données, les connexions aux serveurs de fichiers ou les connexions Cloud.	L'impact d'un Physical Model affiche les Logical Models qui sont mappés vers le Physical Model. L'impact affiche également les Model Stores contenant le modèle.
	Logical Model	La généalogie d'un Logical Model affiche le Physical Model qui est mappé vers les entités du Logical Model.	L'impact d'un Logical Model affiche également les Model Stores contenant le modèle.
	Entité d'un Physical Model ou d'un Logical Model, ou relation d'un Logical Model	À l'intérieur d'un Physical Model, les entités ne comportent pas de la généalogie. À l'intérieur d'un Logical Model, la généalogie d'une entité est la relation entre l'entité et d'autres entités.	À l'intérieur d'un Physical Model, les entités n'ont pas d'impact. À l'intérieur d'un Logical Model, les relations ont un impact sur les entités qu'elles connectent.

Utilise

Modification d'une table de base de données


Lorsque vous modifiez une base de données, cette modification peut avoir des conséquences inattendues sur les systèmes qui s'appuient sur ses données. Les modifications des schémas, des clés, des colonnes ou des tables peuvent affecter les processus en aval et, au final, les rapports, de manières parfois évidentes et parfois subtiles. C'est pourquoi, si vous envisagez de modifier une base de données, vous voulez être sûr de bien comprendre l'impact de vos modifications en aval, de sorte que vous puissiez modifier d'autres systèmes, si nécessaire, pour garantir que vos utilisateurs commerciaux continuent à disposer de données précises et fiables en fonction desquelles prendre les décisions commerciales.

Spectrum™ Technology Platform Metadata Insights fournit une représentation visuelle du flux de données au sein de vos flux, bases de données et fichiers. Vous pouvez sélectionner la base de données que vous modifiez et afficher les flux, les bases de données et les fichiers qui utilisent les données. Grâce à ces informations, vous pouvez déterminer l'impact des modifications de la base de données. Pour plus d'informations, reportez-vous à la section [Affichage de la généalogie et de l'analyse de l'impact](#) à la page 133.

1. Dans un navigateur Web, accédez à :

```
http://server:port/metadata-insights
```

Où : *server* est l'adresse IP ou le nom de serveur de votre serveur Spectrum™ Technology Platform et *port* le port HTTP. Par défaut, le port HTTP est 8080.

2. Cliquez sur le bouton Sélectionner une entité .
3. Cliquez sur **Tables de base de données**.
4. Sélectionnez la table que vous souhaitez modifier, puis cliquez sur **OK**.

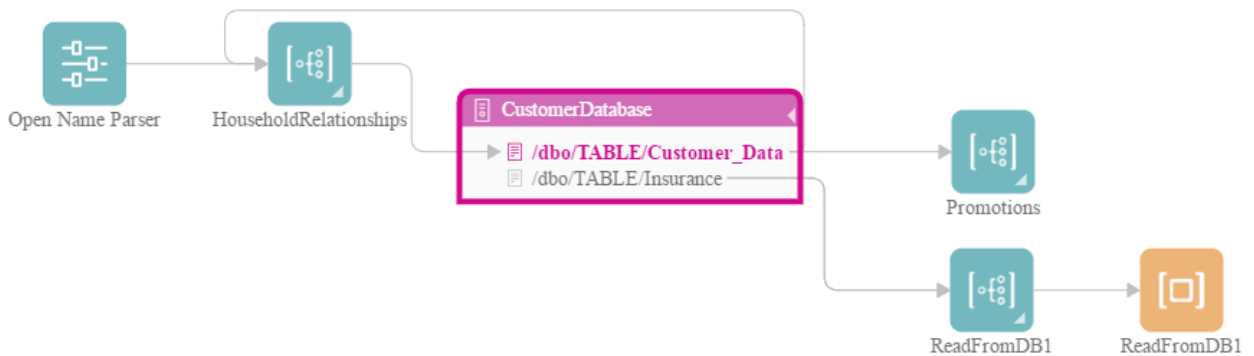
Le diagramme qui en résulte affiche la connexion contenant la table. À gauche de la connexion (généalogie de la table) figure une entité qui écrit des données dans la table. À droite de la table figurent toutes les entités qui utilisent les données de la table, telles que les flux.

Exemple

Vous travaillez sur un projet de standardisation des noms de colonne de plusieurs bases de données. Vous avez trouvé une colonne de la table Customer_Data que vous souhaitez renommer. Avant de modifier le nom de colonne, vous souhaitez connaître les entités qui utilisent les données de cette table, afin de pouvoir les

analyser pour savoir si elles doivent être mises à jour pour prendre en compte le nouveau nom de colonne.

Pour ce faire, vous ouvrez Metadata Insights, vous accédez à Analyse de généalogie et d'impact et vous ouvrez le navigateur d'entités. Vous cliquez sur **Connexions**, puis sur le nom de la connexion à la base de données contenant la table. Ensuite, vous sélectionnez la table, puis vous cliquez sur **OK**. Dans le diagramme qui en résulte, vous développez l'icône de connexion contenant la table, qui vous indique les flux connectés à la table :



Le diagramme révèle que le flux `HouseholdRelationships` écrit des données dans la table `Customer_Data` et lit également ses données. De plus, le flux `Promotions` lit les données de la table `Customer_Data`. Grâce à la consultation de ce diagramme, vous pouvez ouvrir les flux `HouseholdRelationships` et `Promotions` dans Enterprise Designer, pour savoir s'il faut apporter des modifications pour permettre le renommage des colonnes de la table `Customer_Data`.

Mise à jour d'une base de données Spectrum

Les bases de données Spectrum contiennent des données de référence que certains modules utilisent pour valider ou améliorer vos données. Parmi les exemples de bases de données Spectrum figurent la base de données postales utilisée par le module Universal Addressing et les bases de données de géocodage utilisées par les modules Geocoding. Les bases de données Spectrum doivent être régulièrement mises à jour pour vous assurer que vous disposez des toutes dernières données de référence. Lorsqu'il est temps de mettre à jour une base de données Spectrum, vous voulez être sûr que la mise à jour se passera sans encombre et sans interrompre vos opérations.

Pour procéder à une mise à jour sans encombre, vous devez connaître les processus qui utilisent l'ancienne base de données. Si vous connaissez les processus qui utilisent l'ancienne base de données, vous pouvez :


- Les mettre à jour pour qu'ils puissent utiliser la nouvelle base de données.
- Les tester avec la nouvelle base de données pour vous assurer d'obtenir les résultats escomptés.

L'outil Généalogie et Analyse de l'impact fournit la visibilité dont vous avez besoin pour effectuer une mise à jour sans encombre d'une base de données Spectrum. Il présente les services et les flux qui utilisent une base de données Spectrum, de sorte que vous puissiez les mettre à jour pour qu'ils pointent vers la nouvelle base de données et les tester pour vous assurer qu'ils continuent à produire les résultats escomptés après la mise à jour.

1. Dans un navigateur Web, accédez à :

`http://server:port/metadata-insights`

Où : *server* est l'adresse IP ou le nom de serveur de votre serveur Spectrum™ Technology Platform et *port* le port HTTP. Par défaut, le port HTTP est 8080.

2. Cliquez sur le bouton Sélectionner une entité .
3. Cliquez sur **Bases de données Spectrum**.
4. Cliquez sur le bouton en regard du nom du module pour le sélectionner, puis cliquez sur **OK**. Toutes les bases de données Spectrum du module sélectionné apparaissent.

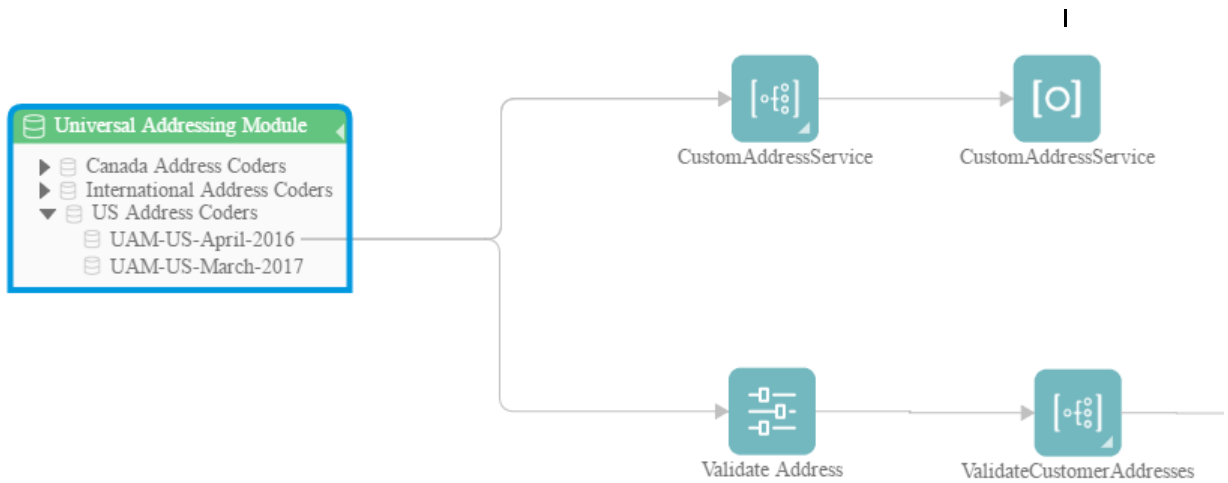
Conseil : Même si vous pouvez cliquer sur le nom de module lui-même pour afficher les bases de données individuelles qui sont utilisées par le module, nous vous recommandons de commencer par afficher toutes les bases de données du module, pour que vous ayez une vue complète de toutes les bases de données Spectrum utilisées pour le module.

Le diagramme qui en résulte affiche les entités qui s'appuient sur chaque base de données Spectrum pour le module sélectionné.

Exemple

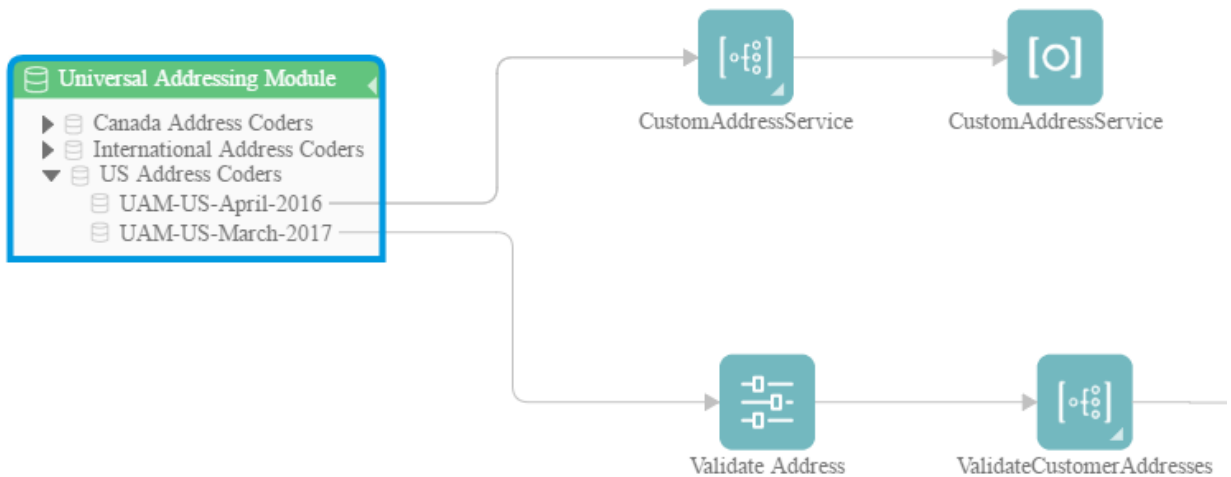
Vous utilisez le module Universal Addressing pour valider les adresses de votre base de données clients. Vous utilisez actuellement une base de données postales obsolète et souhaitez mettre votre système à jour pour pouvoir utiliser la toute dernière base de données. Vous avez installé la base de données et lui avez défini une base de données Spectrum dans Management Console. Maintenant, vous devez savoir où l'ancienne base de données est utilisée, afin de pouvoir mettre à jour tous les composants pour qu'ils utilisent la base de données, et afin de tester tous les processus affectés pour garantir qu'ils produisent les résultats escomptés avec la nouvelle base de données.

Ouvrez Metadata Insights, accédez à Généalogie et Analyse de l'impact et ouvrez le navigateur d'entités. Cliquez sur **Bases de données Spectrum**. Étant donné que vous souhaitez afficher toutes les ressources de base de données du module Universal Addressing, sélectionnez le bouton en regard de **Module Universal Addressing**, puis cliquez sur **OK**. Le diagramme qui en résulte est le suivant :

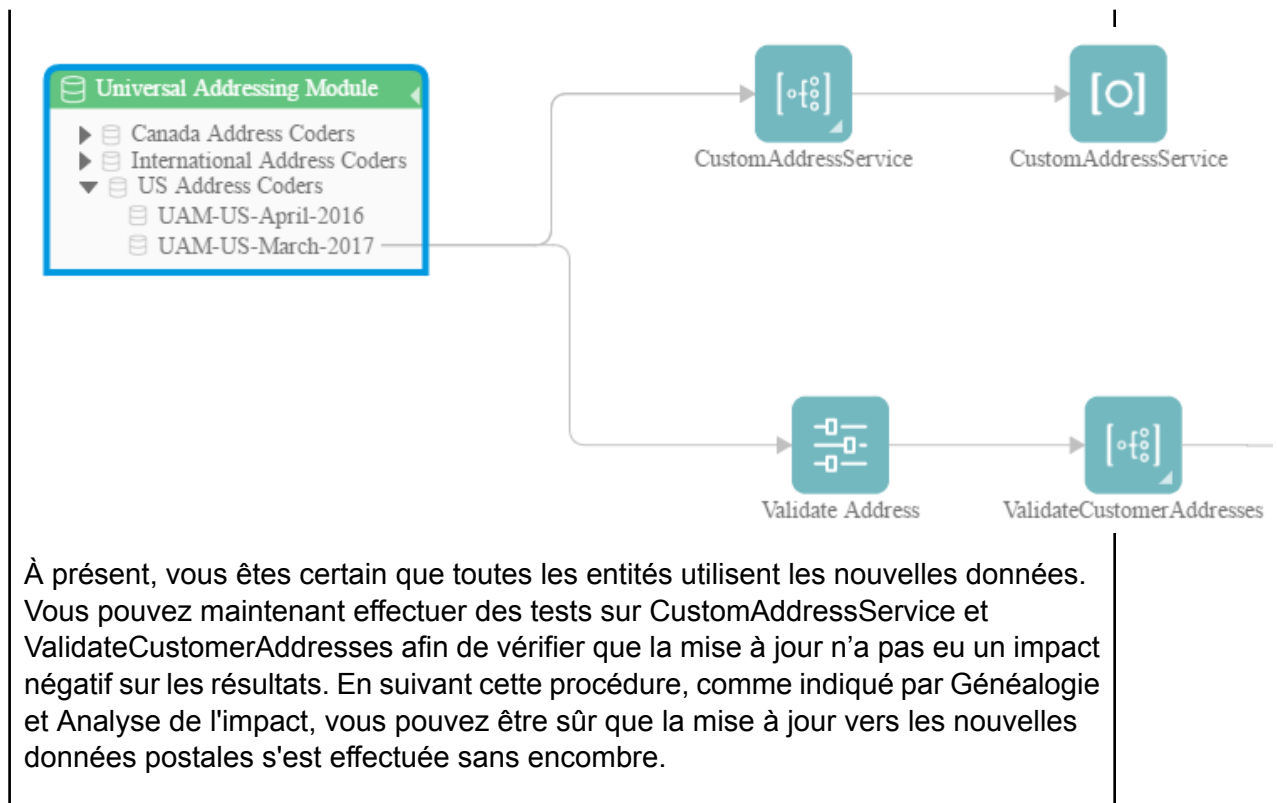


Le diagramme indique qu'il existe deux bases de données Spectrum sur le système, l'ancienne (UAM-US-April-2016) et la nouvelle que vous venez juste de créer dans Management Console (UAM-US-March-2017). Il existe deux entités qui utilisent l'ancienne base de données : un flux nommé CustomAddressService et un flux nommé ValidateCustomerAddresses. Il n'existe aucune entité qui utilise la nouvelle base de données.

Mettez à jour les options de service par défaut de sorte que Validate Address utilise la nouvelle base de données. Après l'actualisation du diagramme, vous voyez apparaître ceci :



Comme vous pouvez le constater, les options de service par défaut de Validate Address utilisent maintenant la nouvelle base de données (UAM-US-March-2017), mais le flux CustomAddressService continue à utiliser l'ancienne base de données (UAM-US-April-2016). Ouvrez le flux CustomAddressService dans Enterprise Designer et modifiez le stage Validate Address afin qu'il utilise la nouvelle base de données. Après l'actualisation du diagramme, vous voyez apparaître ceci :



Modification des options par défaut d'un service


Les services Spectrum™ Technology Platform comportent des options par défaut que vous configurez dans Management Console. Les options par défaut prennent effet lorsqu'un service est utilisé sous forme de stage dans un flux et que le stage est configuré pour utiliser les options par défaut du système. Par conséquent, avant d'apporter des modifications aux options du service dans Management Console, vous devez savoir quels flux utilisent les options par défaut du service, pour vous assurer que la modification n'a pas de conséquences indésirables sur ces flux ni sur aucun des processus commerciaux qui s'appuient sur ces flux.

Pour trouver les flux qui utilisent les options par défaut d'un service :

1. Dans un navigateur Web, accédez à :

`http://server:port/metadata-insights`

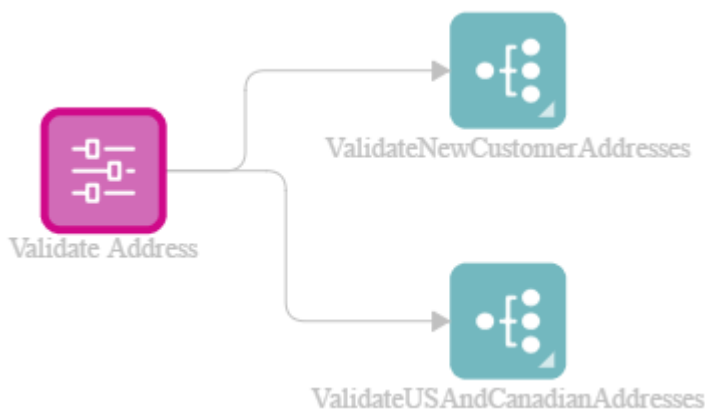
Où : *server* est l'adresse IP ou le nom de serveur de votre serveur Spectrum™ Technology Platform et *port* le port HTTP. Par défaut, le port HTTP est 8080.

2. Cliquez sur le bouton Sélectionner une entité .
3. Cliquez sur **Options de service par défaut**.
4. Sélectionnez le service dont vous souhaitez modifier l'option par défaut, puis cliquez sur **OK**.

Le diagramme qui en résulte affiche tous les flux qui utilisent les options par défaut du service. En d'autres termes, il s'agit des flux dans lesquels l'option de stage **Neutralisation des options système par défaut avec les valeurs suivantes** n'est pas cochée. Si un flux contient le service, mais que cette option est cochée, il n'est pas inclus dans le diagramme, car il n'utilise pas les options par défaut du service et, par conséquent, ne serait pas affecté par les modifications apportées aux options par défaut du service dans Management Console.

Exemple

Vous avez configuré le service Validate Address pour qu'il renvoie les adresses en majuscules et vous souhaitez remplacer l'option de casse par une casse mixte dans Management Console, ce qui signifie que tous les stages Validate Address qui utilisent les options par défaut du service vont commencer à renvoyer des adresses en casse mixte. Pour vous assurer que cette modification n'a pas de conséquences indésirables sur ces flux, utilisez la vue Généalogie et Analyse de l'impact de Metadata Insights pour savoir quels flux utilisent les options par défaut de Validate Address. Elle affiche ce diagramme :



Le diagramme montre que les flux de ValidateNewCustomerAddresses et ValidateUSAndCanadianAddresses utilisent tous les deux les options par défaut du service Validate Address. Étant donné que toutes les modifications apportées aux options de Validate Address dans Management Console affecteront ces deux flux, ouvrez-les dans Enterprise Designer pour examiner l'impact de la modification sur la manière dont les flux gèrent les adresses. Si vous pensez que le passage à une casse mixte entraîne des problèmes, vous pouvez ajuster le flux de sorte qu'il prenne en charge la modification.

Modification d'un sous-flux

Un sous-flux est un flux qui peut être réutilisé au sein d'autres flux. Les sous-flux s'avèrent utiles si vous souhaitez créer un processus réutilisable, facile à intégrer dans différents flux. Lorsque les


sous-flux sont beaucoup utilisés, il peut être difficile de savoir quels flux utilisent chaque sous-flux. Cela peut poser problème si vous souhaitez modifier un sous-flux, car vous devez vous assurer que la modification n'entraîne pas de résultats indésirables dans les flux qui l'utilisent. À l'aide de la fonction Généalogie et Analyse de l'impact de Metadata Insights, vous pouvez repérer tous les flux qui utilisent un sous-flux donné, pour être parfaitement informé de l'impact des modifications que vous apportez à un sous-flux.

Pour afficher les flux qui utilisent un sous-flux :

1. Dans un navigateur Web, accédez à :

`http://server:port/metadata-insights`

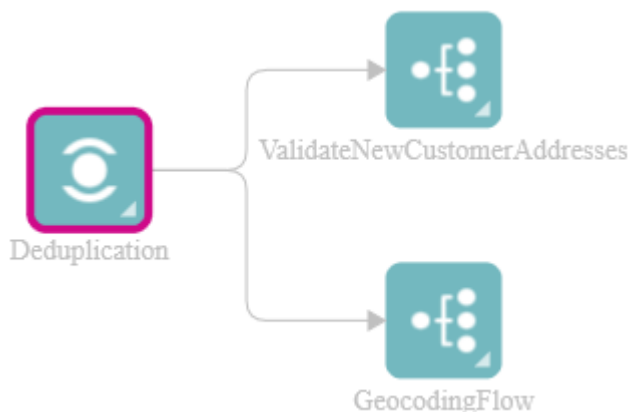
Où : *server* est l'adresse IP ou le nom de serveur de votre serveur Spectrum™ Technology Platform et *port* le port HTTP. Par défaut, le port HTTP est 8080.

2. Cliquez sur le bouton Sélectionner une entité .
3. Cliquez sur **Stages**.
4. Sélectionnez le sous-flux de votre choix, puis cliquez sur **OK**.

Le diagramme qui en résulte affiche tous les flux qui utilisent le sous-flux.

Exemple

Vous disposez d'un sous-flux nommé Déduplication. Il contient une logique permettant d'identifier les enregistrements en double dans vos données client. Vous utilisez ce sous-flux dans différents flux. Vous souhaitez modifier la logique de déduplication et devez vous assurer de tester les effets de la modification dans tous les flux qui utilisent le sous-flux. Dans ce cas, utilisez Généalogie et Analyse de l'impact pour voir l'impact du sous-flux Déduplication :



Le diagramme Généalogie et Analyse de l'impact indique que le sous-flux Déduplication est utilisé dans deux flux : ValidateNewCustomerAddresses et GeocodingFlow. Lorsque vous apportez votre modification, vous pouvez vérifier ces flux dans Enterprise Designer pour confirmer que vous obtenez les résultats escomptés.

Notices

© 2017 Pitney Bowes Software Inc. Tous droits réservés. MapInfo et Group 1 Software sont des marques commerciales de Pitney Bowes Software Inc. Toutes les autres marques et marques commerciales sont la propriété de leurs détenteurs respectifs.

Avis USPS®

Pitney Bowes Inc. détient une licence non exclusive pour la publication et la vente de bases de données ZIP + 4® sur des supports optiques et magnétiques. Les marques de commerce suivantes appartiennent à United States Postal Service : CASS, CASS Certified, DPV, eLOT, FASTforward, First-Class Mail, Intelligent Mail, LACS^{Link}, NCOA^{Link}, PAVE, PLANET Code, Postal Service, POSTNET, Post Office, RDI, Suite^{Link}, United States Postal Service, Standard Mail, United States Post Office, USPS, ZIP Code et ZIP + 4. Cette liste de marques de commerce appartenant à U.S. Postal Service n'est pas exhaustive.

Pitney Bowes Inc. détient une licence non exclusive de USPS® pour le traitement NCOA^{Link®}.

Les prix des produits, des options et des services de Pitney Bowes Software ne sont pas établis, contrôlés ni approuvés par USPS® ni par le gouvernement des États-Unis. Lors de l'utilisation de données RDI™ pour déterminer les frais d'expédition de colis, le choix commercial de l'entreprise de distribution de colis à utiliser n'est pas fait par USPS® ni par le gouvernement des États-Unis.

Fournisseur de données et avis associés

Les produits de données contenus sur ce support et utilisés au sein des applications Pitney Bowes Software sont protégés par différentes marques de commerce et par un ou plusieurs des copyrights suivants :

© Copyright United States Postal Service. Tous droits réservés.

© 2014 TomTom. Tous droits réservés. TomTom et le logo TomTom logo sont des marques déposées de TomTom N.V.

© 2016 HERE

Source : INEGI (Instituto Nacional de Estadística y Geografía)

Basées sur les données électroniques © National Land Survey Sweden.

© Copyright United States Census Bureau

© Copyright Nova Marketing Group, Inc.

Des portions de ce programme sont sous © Copyright 1993-2007 de Nova Marketing Group Inc. Tous droits réservés.

© Copyright Second Decimal, LLC

© Copyright Canada Post Corporation

Ce CD-ROM contient des données provenant d'une compilation dont Canada Post Corporation possède le copyright.

© 2007 Claritas, Inc.

Le jeu de données Geocode Address World contient des données distribuées sous licence de GeoNames Project (www.geonames.org) fournies sous la licence Creative Commons Attribution License (« Attribution License ») à l'adresse :

<http://creativecommons.org/licenses/by/3.0/legalcode>. Votre utilisation des données GeoNames (décrites dans le Manuel de l'utilisateur Spectrum™ Technology Platform) est régie par les conditions de la licence Attribution License et tout conflit entre votre accord avec Pitney Bowes Software, Inc. et la licence Attribution License sera résolu en faveur de la licence Attribution License uniquement s'il concerne votre utilisation des données GeoNames.



3001 Summer Street
Stamford CT 06926-0700
USA

www.pitneybowes.com