

Catalogage intelligent des données pour les data warehouses, data lakes et lakehouses dans le Cloud

Principaux avantages

- Optimisez la valeur des data warehouses, des data lakes et des lakehouses dans le Cloud sans interruption
- Accélérez les initiatives de modernisation et de migration vers le Cloud, même les plus complexes
- Gagnez en visibilité sur toutes les ressources de données de votre entreprise grâce à une gestion des métadonnées de pointe
- Facilitez la recherche de données pertinentes dans le Cloud pour les utilisateurs

Accélérer l'intégration de données dans le Cloud et la découverte de données basée sur l'IA

Aujourd'hui, les entreprises transfèrent rapidement leur infrastructure de données et d'analyse vers le Cloud. Si cette volonté de capitaliser sur l'agilité, la scalabilité et la rentabilité du Cloud affecte tous les aspects de l'analyse des données, elle est particulièrement d'actualité en ce qui concerne le data warehousing et les data lakes. Pendant des années, les data warehouses d'entreprise ont été des outils de travail stables, qui ont alimenté les systèmes d'analyse et de génération de rapports d'entreprise. L'évolution actuelle vise à les moderniser dans le Cloud afin d'améliorer considérablement les performances et l'avantage concurrentiel. Dans le même temps, les entreprises commencent à utiliser des data lakes dans le Cloud pour aller de pair avec leurs data warehouses dans le Cloud. Et plus récemment, les entreprises ont commencé à construire des lakehouses dans le Cloud, qui associent les avantages du data warehouse et du data lake pour fournir des capacités d'analyse permettant d'alimenter les tableaux de bord BI comme les projets avancés d'IA et de machine learning.

Toutefois, l'adoption de data warehouses, de data lakes et de lakehouses dans le Cloud peut présenter de nouveaux défis. Même s'il est facile de mettre en place un nouveau data warehouse ou data lake dans le Cloud, vous avez besoin d'une stratégie et d'une certaine planification pour optimiser la valeur de votre investissement.

Que vous construisiez un nouveau data warehouse, data lake ou lakehouse dans le Cloud, ou bien que vous modernisiez les données et les charges de travail dans le Cloud au fil du temps, vous devez absolument comprendre et évaluer votre environnement de données actuel. Assurez-vous aussi que vous disposez des outils et des meilleures pratiques nécessaires pour gérer vos données une fois celles-ci dans le Cloud.

La solution de gestion intelligente des données de lakehouse dans le Cloud pour Cloud Analytics d'Informatica® repose sur les capacités les plus performantes du marché en matière d'intégration de données, de qualité des données et de gouvernance des données, ainsi que de gestion des métadonnées. Informatica Enterprise Data Catalog fournit la base de gestion des métadonnées qui vous permet de comprendre votre paysage de données grâce à la découverte des ressources de données basée sur l'apprentissage automatique et à une visibilité de bout en bout. Informatica Intelligent Cloud ServicesSM (IICS) fournit des services iPaaS d'entreprise leaders sur le marché dédiés à l'intégration de données dans le Cloud, l'ingestion en masse dans le Cloud et la qualité des données dans le Cloud.

Grâce à ses fonctions de machine learning et à sa gestion complète des métadonnées, Informatica Enterprise Data Catalog fournit une base commune de métadonnées d'entreprise pour une visibilité et une traçabilité intelligentes, automatisées et de bout en bout dans votre environnement. Cette solution vous permet de relier les magasins de données sur site à votre écosystème Cloud moderne pour optimiser la valeur sans perturber les consommateurs de données.

Fonctionnalités clés

Connectivité étendue des métadonnées

La solution de catalogage intelligent des données d'Informatica s'intègre à tous les écosystèmes Cloud de pointe et aux data warehouses, data lakes et lakehouses modernes, notamment Amazon Redshift, Amazon S3, Microsoft Azure Synapse Analytics, Google BigQuery, Snowflake et Databricks. Elle inclut également une connectivité des métadonnées prête à l'emploi pour un accès à des centaines d'applications Cloud, de systèmes sur site, d'applications d'entreprise et middleware, de magasins de données et d'outils BI analytiques.

Découverte des données basée sur le machine learning

CLAIRE[®], le moteur d'intelligence artificielle d'Informatica basé sur les métadonnées, applique des algorithmes d'IA et de machine learning aux données techniques, commerciales, opérationnelles et d'utilisation. Il comprend et analyse automatiquement les fichiers et métadonnées complexes, vous permettant ainsi de découvrir les données où qu'elles se trouvent dans un environnement hybride et multi-Cloud.

Traçabilité des données automatisée de bout en bout

La traçabilité vous offre des vues de synthèse et des vues détaillées qui vous permettent de connaître l'origine de vos données, leur déplacement au sein de votre entreprise ainsi que leur transformation. La traçabilité des données de bout en bout dans des environnements hybrides et multi-Cloud vous permet de suivre les données de la source à la cible. Vous savez ainsi comment les données sont utilisées et ce qui en dépend. Avec Informatica Enterprise Data Catalog, vous pouvez effectuer une analyse détaillée de l'impact sur les ressources de données en amont comme en aval. Ainsi, vous comprenez facilement l'impact de la migration sur les ressources de données et autres ressources, ainsi que sur les utilisateurs de données. En outre, grâce à la traçabilité détaillée des données, l'analyse de l'impact peut vous aider à démontrer les avantages financiers liés au transfert de certaines ressources de données et charges de travail vers le Cloud.

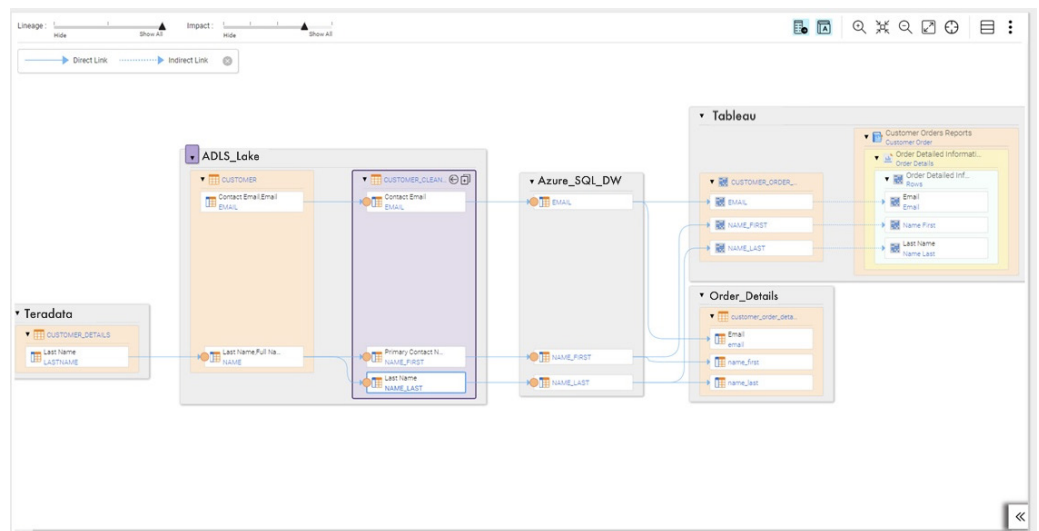


Figure 1. Visualisez facilement la traçabilité des données de bout en bout sur site ou dans les data warehouses, data lakes et lakehouses dans le Cloud.

Collaboration et curation sociale

La migration de données vers un data warehouse, data lake ou lakehouse dans le Cloud implique généralement plusieurs intervenants et utilisateurs. Informatica Enterprise Data Catalog propose des fonctionnalités de collaboration et de curation sociale qui permettent aux propriétaires de données et aux experts en la matière de certifier les ressources de données et d'identifier celles qui doivent être migrées. De plus, les consommateurs de données peuvent évaluer et examiner les ressources de données, en ajoutant des attributs et des annotations personnalisés qui facilitent la recherche de données fiables pendant et après la migration.

Intégration aux services de gestion des données dans le Cloud

Informatica Enterprise Data Catalog s'intègre aux services de gestion de données IICS qui fournissent des fonctionnalités critiques pour les data warehouses, data lakes et lakehouses dans le Cloud. IICS Cloud Data Integration offre des fonctionnalités ETL, d'ingestion de données, de synchronisation et de réplication hautes performances pour les environnements multi-Cloud. IICS Cloud Mass Ingestion prend en charge l'ingestion de flux, de fichiers et de bases de données dans les data warehouses et data lakes dans le Cloud, ainsi que les files d'attente de messages. IICS Cloud Data Quality vous permet de profiler, de nettoyer et de normaliser les données afin de surveiller et de résoudre rapidement les problèmes de qualité des données. Dans IICS, vous pouvez utiliser Informatica Enterprise Data Catalog pour découvrir et sélectionner des données, afficher les sources et les mappings de données, afficher les règles de qualité de données et les fiches d'évaluation, et bénéficier d'une visibilité complète sur la traçabilité des données de bout en bout.

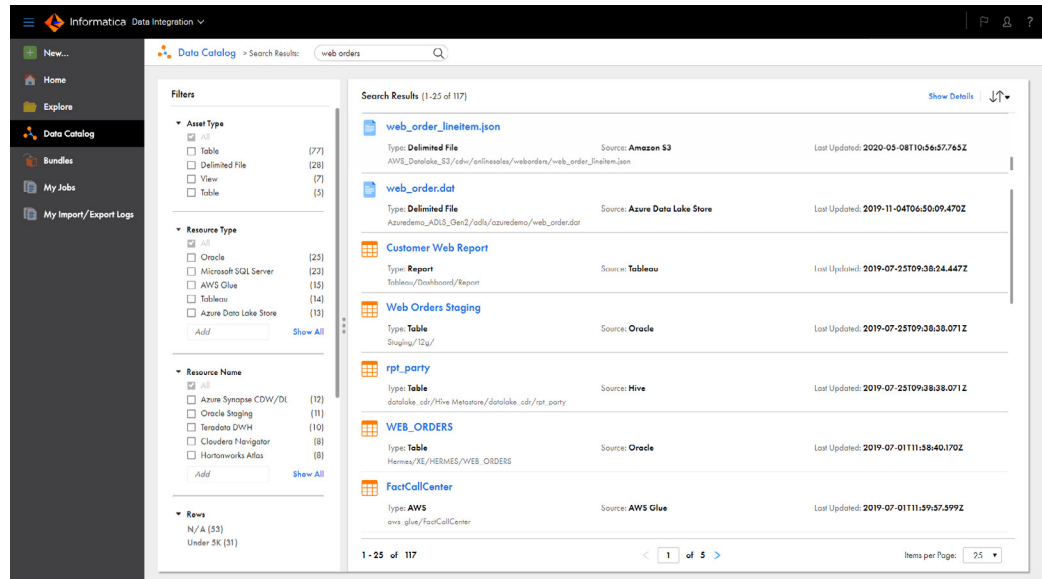


Figure 2. Découvrez et sélectionnez les données à l'aide d'Informatica Enterprise Data Catalog dans IICS.

Mappings préconfigurés

Les modèles de mapping préconfigurés, les assistants de tâches faciles à utiliser pour la réplication et la synchronisation des données, ainsi que l'ingestion de données en masse pour les charges de travail volumineuses vous permettent de déplacer davantage de données vers le Cloud en moins de temps. L'intégration transparente entre Informatica Enterprise Data Catalog et IICS vous permet de trouver les données à migrer et de les déplacer facilement vers votre data warehouse, data lake ou lakehouse dans le Cloud via une seule interface sans code. « Achetez » votre catalogue de données avec le concepteur visuel d'IICS Cloud Data Integration, puis créez des charges de données simples à complexes en quelques clics pour connecter différents terminaux sur le Cloud et sur site, transformer les données et planifier les charges de travail d'analyse.

Principaux avantages

Optimisez la valeur des data warehouses, des data lakes et des lakehouses dans le Cloud sans interruption

Évaluez de manière exhaustive votre paysage de données afin de comprendre quoi migrer ou consolider, et où vous allez générer le plus de valeur à partir d'un data warehouse, data lake ou lakehouse dans le Cloud. Grâce à la traçabilité des données de bout en bout, vous gagnez en visibilité sur les relations et dépendances entre les données. Offrez aux consommateurs de données un accès en libre-service pour trouver facilement les données à des fins d'analyse dans le Cloud. Informatica Enterprise Data Catalog vous garantit que les structures et les schémas de votre data warehouse sur site sont transférés vers le Cloud, évitant ainsi toute interruption potentielle pendant et après la migration. Après la migration, les vues de traçabilité de bout en bout vous aident à confirmer que les données ont été mappées correctement. En outre, la curation et la découverte de données automatisées basées sur les métadonnées prennent en charge la gouvernance évolutive des données et permettent la livraison de données fiables dans le Cloud.

Accélérez les initiatives de modernisation et de migration vers le Cloud, même les plus complexes

Réduisez considérablement le temps nécessaire à la modernisation et à la migration vers le Cloud. Découvrez les données à migrer ou à consolider, puis créez des charges d'intégration de données simples ou complexes à l'aide d'un concepteur visuel de mappings qui connecte divers terminaux sur le Cloud et sur site avec des transformations prédéfinies pour accélérer le développement.

Gagnez en visibilité sur toutes les ressources de données de votre entreprise grâce à une gestion des métadonnées de pointe

Vous devez découvrir et comprendre vos données pour en tirer de la valeur. La solution d'Informatica s'appuie sur les meilleures fonctionnalités de gestion des métadonnées du marché, mettant ainsi à disposition des centaines de connecteurs pour les systèmes basés sur le Cloud et sur site. Grâce à l'intelligence basée sur les métadonnées, cette solution vous aide à identifier tous les endroits où les données sont stockées sur l'ensemble des systèmes de votre entreprise. Vous pouvez ensuite déplacer toutes ces données de manière transparente vers le data warehouse, data lake ou lakehouse dans le Cloud de votre entreprise. En outre, cette solution capture et stocke les métadonnées critiques concernant les imprévus et les relations entre les ensembles de données et les sous-ensembles, afin de maintenir les connexions entre eux pendant et après la modernisation et la migration des données vers le Cloud.

À propos d'Informatica

La transformation digitale fait évoluer les attentes : meilleurs services, livraisons plus rapides, à moindre coût. Les données sont la clé de la réussite des entreprises, et ces dernières doivent évoluer pour rester compétitives.

En tant que leader mondial dans la gestion des données Cloud d'entreprise, nous sommes prêts à vous guider de manière intelligente – quel que soit le secteur, la catégorie ou la niche. Informatica vous permet de prendre une longueur d'avance pour gagner en agilité, concrétiser de nouvelles opportunités de croissance ou même innover. Nous nous concentrons sur les données afin de vous offrir la polyvalence nécessaire pour réussir.

Découvrez nos solutions et libérez tout le potentiel de vos données en vue de la prochaine révolution intelligente.

Facilitez la recherche de données pertinentes et fiables dans le Cloud pour les utilisateurs

Lorsque des données sont déplacées au cours d'une migration, les consommateurs de données peuvent avoir des difficultés à localiser les données dont ils dépendent pour réaliser leur travail. Il est possible que les liens marqués d'un signet et les rapports récurrents se perdent. Les consommateurs de données peuvent vouloir continuer à utiliser des données provenant d'une source sur site familière, alors que le Cloud constitue désormais une meilleure solution. Les analystes et les ingénieurs de données ont également besoin de savoir où se trouvent les données et de comprendre quelles ressources sont les mieux adaptées à leurs besoins. Grâce à Informatica Enterprise Data Catalog, les utilisateurs trouvent facilement les données dont ils ont besoin sans avoir à demander d'aide au service informatique, ce qui évite les goulets d'étranglement et les inefficacités.

Étapes suivantes

Pour en savoir plus, consultez la [fiche technique sur Enterprise Data Catalog d'Informatica](#) et la présentation de solution [Modernisez votre data warehouse et votre data lake dans le Cloud](#).

