



Mit ETL-Datenströmen Geschäftswissen erfolgreich generieren

Effektive Datenintegration mit Talend Open Studio



Die Befüllung von Data Warehouses ist heute wichtiger Bestandteil der BI-Architektur in den Unternehmen. Die Kosten für solche ETL-Entwicklungsumgebungen spielen eine wichtige Rolle. Deshalb suchen Firmen oft nach effizienten, leistungsfähigen und vollwertigen Lösungen, die sich auch im Open-Source-Umfeld finden lassen. Mit Talend Open Studio for Data Integration steht hier eine effiziente Lösung zur Modellierung und Umsetzung von Datenmigrations-, Datensynchronisations- und Datenintegrationsprozesse zur Verfügung.

Talend bietet Open-Source-Software für Datenintegration, Datenmanagement, Enterprise Application Integration und Big-Data-Software auf dem Markt an. Subscription-Modelle ergänzen die Produkte mit zusätzlichen-Funktionalitäten und sichern Support-Leistungen, die mit den Leistungen rein kommerzieller Anbieter vergleichbar sind.

Talend's Datenintegrationsprodukte stellen eine erweiterbare Palette von Open-Source-Tools bereit, mit der sich Daten aus jedem Business-System in Echtzeit oder im Batch-Modus umwandeln und integrieren lassen. Dabei werden operative und analytische Anforderungen der Datenintegration bis hin zu Big-Data-Anbindungen berücksichtigt.

Was bietet der Data Integrator?

Um die Entwicklung von ETL-Prozesse zu unterstützen, bringt Talend Open Studio for Data Integration eine grafische Entwicklungsumgebung auf Grundlage von Eclipse mit. Die Entwicklungsumgebung lässt sich an Unternehmensanforderungen individuell anpassen und teilt sich in drei Hauptanwendungen auf:

Business Modeler

Zur Darstellung von Geschäftsprozessen hält Talend Data Integrator den Business Modeler bereit. Hiermit können Unternehmen ihre Geschäftsprozesse leicht skizzieren und mögliche Probleme und Schwierigkeiten vorab sichtbar machen. Das visuelle Werkzeug gruppiert alle relevanten Unterlagen zur Datenintegration in einem Geschäftsplan.

Job Designer

Bei der Erstellung von ETL-Prozessen ist es wichtig, den Überblick schnell zu erhalten und immer das Wesentliche zu sehen. Auf diese Weise lassen sich bei der Entwicklungsphase unnötige Fehler vermeiden. Auch bei der Weiterentwicklung und Wartung der Prozesse ist es wichtig, schnell zu erkennen, welche Möglichkeiten oder Verbesserungspotenziale bestehen.

Diese Aufgabe übernimmt der Job Designer. Dieser bietet sowohl eine grafische als auch eine funktionale Sicht auf die tatsächlichen Integrationsprozesse. Über eine Palette von Komponenten und Konnektoren lassen sich durch einfaches Ziehen und Ablegen auf dem Arbeitsbereich umfangreiche Integrationsprozesse gestalten. Komplexe Prozesse lassen sich mit den intuitiven grafischen Komponenten und dem integrierten Assistenten effektiv und einfach erstellen. Zur höheren Lesbarkeit eines Arbeitsbereiches lassen sich die Prozesse in Teilbereiche zerlegen. Die Eigenschaften werden unter den Komponenten eingebunden und können mittels Metadata Manager wiederverwendet und vererbt werden.

Metadata Manager

Sämtliche Projektinformationen und Unternehmensmetadaten werden in einem zentralen Repository konsolidiert, das alle Stakeholder, also Business-Anwender, Job-Entwickler, IT-Mitarbeiter, gemeinsam nutzen können. Die Verwaltung des Repository wird über den Metadata Manager vorgenommen. Das Repository zentralisiert alle Projektinformationen und gewährleistet in allen Integrationsprozessen die notwendige Konsistenz. Kontextdaten wie Datenbankverbindung und Dateipfade können hier zentralisiert werden. Dies erleichtert die Aktualisierung der Unternehmensanwendungen. Auch lassen sich Routinen, Methoden und Teile von Programmierungen in dieser Anwendung vereinheitlichen.

Test und Debugging in Echtzeit

Um Fehler bei der Entwicklung und unnötiges Zeitaufkommen bei der Fehleranalyse zu vermeiden, sind leistungsstarke Test- und Debugging-Werkzeuge notwendig. Talend gibt in diesem Bereich diverse Werkzeuge an die Hand, die beispielsweise das Echtzeit-Tracking von Daten über den gesamten Transformationsprozess ermöglichen. Mit diesen lässt sich das Geschehen zur Laufzeit sichtbar machen.

Schnittstellen- und Formatvielfalt

Talend stellt über 550 native Verbindungen zu Datenbanken, Anwendungspaketen wie ERP, CRM etc., SaaS- und Cloud-Anwendungen, Mainframes, Dateien, Webdiensten, Data Warehouses, Data Marts und OLAP-Anwendungen her. Es stehen integrierte und erweiterte Komponenten für ETL inklusive String-Manipulatoren, Slowly Changing Dimensions, automatischem Lookup-Handling und Bulkloading zur Verfügung. Die direkte Integration ist u. a. mit Funktionen zur Datenqualität, zum Datenabgleich und zur Stammdatenverwaltung ausgestattet. Talend bietet eine einfache Konnektivität zu den gängigsten Cloud-Apps wie Salesforce.com und SugarCRM.

Open Source und Enterprise Edition

Talend Open Studio for Data Integration ist als Open Source oder in verschiedenen Enterprise Editionen erhältlich. Mit allen Enterprise Editionen erhält man zusätzlich das Talend Administration Center. Hiermit lassen sich die Benutzer, Ausführungspläne für definierte Jobs und die Integrationsjobs verwalten.

Ergänzende Produkte

Talend stellt noch diverse andere Produkte zu Verfügung, die die ETL Prozesslandschaft in einem Unternehmen vervollständigen und abrunden. Hier ein Ausschnitt aus dem Produktportfolio:

Big Data

Für die großen Datenmengen steht dem User die Big Data Edition zur Verfügung. Mit dieser Version wird zusätzlich auch die Generierung von Map-Reduce-Jobs mittels Hadoop ermöglicht. Die Version Talend Open Studio for Big Data ist auch als Open-Source-Variante verfügbar.

Master Data Management

Talend bietet eine professionelle Lösung für Master Data Management (MDM) an. Mithilfe von Master Data Management können Stammdaten ihres Unternehmens gepflegt und verwaltet werden. Talend verfügt über eine Enterprise- und eine Community-Edition für MDM.

Data Quality

Talend hat eine Komplettlösung zur Sicherstellung der Unternehmensdatenqualität in seinem Stack. Mit Talend Data Quality lassen sich mitunter inkonsistente Daten eliminieren, Regeln durchsetzen und konsistente Informationen durch Standardisierung erstellen.

Folgen Sie uns



www.opitz-consulting.com/newsroom

Eine große Community hilft weiter

Als Anwender und Entwickler können Sie sich bei Fragen und Problemen Unterstützung in der Online-Community „TalendForge“ holen. Diese verfügt über eine große aktive Gruppe von Usern

Forum: <http://www.talendforge.org/forum/>

Wiki: <http://www.talendforge.org/wiki/>

Bugtracker: <http://www.talendforge.org/bugs/>

Exchange: <http://talendforge.org/exchange/>

Warum machen wir „Werbung“ für Talend?

Talend bietet mit seiner Produktpalette effiziente Enterprise-Integrationslösungen an, die alles Notwendige abdecken, was Sie für Ihr Data-warehouse-Vorhaben benötigen. Nach unserer Erfahrung geht eine Einarbeitung dank der einfachen Handhabung der Werkzeuge zügig vonstatten und der Frustrationsfaktor bleibt dabei gering. Die vorhandene große Anzahl von Verbindungen sichert die Anbindung weiterer Systeme an das Data Warehouse und macht es somit zukunftssicher. Deshalb empfehlen wir unseren Kunden, den Einsatz dieser Werkzeuge zu prüfen.

Talend Gold Partner

Als Talend Gold Partner verfügt unsere IT-Beratung über umfangreiches Know-how und Erfahrungen mit Talend Werkzeugen. Nutzen Sie diese Stärken, um die Tools im BI-Umfeld richtig und professionell einzusetzen! Lassen Sie sich bei der Planung und beim Einsatz in Ihrem Unternehmen von unseren Experten beraten.



Sprechen Sie uns an:

Thomas Eichler

Senior Consultant

Telefon: +49 2261 6001-0

Telefax: +49 2261 6001-4200

E-Mail: thomas.eichler@opitz-consulting.com

Mehr zu unserem Leistungsbereich BI/Analytics:

www.opitz-consulting.com/bi-analytics

