

# Talend Open Studio

# v2

Talend vereinfacht die Datenintegration. Mittels einer offenen und innovativen Plattform bietet Talend eine einfache und leistungsfähige Lösung.

Durch die hohen Kosten von proprietären Lösungen wurden eine Reihe von Unternehmen gezwungen ihre Datenintegrations-Bedürfnisse mittels Eigenentwicklungen und Routinen zu stillen. Meistens jedoch decken diese Integrations-Routinen nur eine ganz spezielle Anforderung ab ohne einen globalen Ansatz oder eine konsolidierte Entwicklung zu fokussieren und ohne Berücksichtigung bereits existierender Prozesse. Dieses Verfahren erzeugt einen hohen Wartungsaufwand, beeinträchtigt die Stabilität, gefährdet die Integrität und Konsistenz von Informationssystemen und spiegelt sich in hohen Wartungskosten wieder.

## Eine vielseitige Datenintegration's Lösung

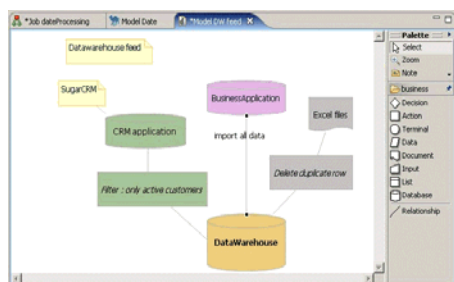
Talend Open Studio ist eine leistungsstarke und vielseitige OpenSource-Lösung welche alle Belange der täglichen Datenintegration abdeckt:

- Synchronisation oder Replikation von Datenbanken
- Pünktlicher Austausch und Verteilung von Daten innerhalb der IT-Umgebung
- ETL (Extraktion, Transformation, Laden) für analytische Systeme
- Datenmigration
- Aufwendiges Daten-Transformieren und Laden

Talend Open Studio besteht aus drei Modulen (Business Modeler, Job Designer und Metadata Manager) welche mittels der einheitlichen graphischen Entwicklungsumgebung, auf Eclipse basierend, sehr einfach zu bedienen sind.

## Geschäftsorientierte Prozessmodellierung

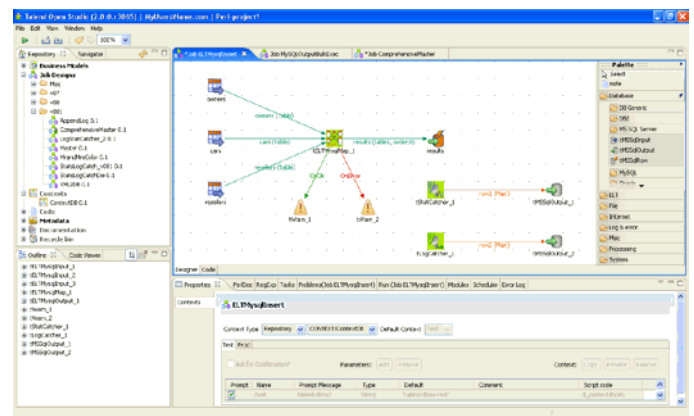
Talend Open Studio's Business Modeler unterstützt den Top-Down Ansatz. Dieses ermöglicht die Einbindung von Geschäftsanwendern in das Design des Integrationsprozesses und zeigt die Fortschritte während der Umsetzung. Der Business Modeler verbindet alle für den Daten-



integrations-Prozess relevanten Dokumentationen in einer anwenderfreundlichen Übersicht. Dieses ermöglicht ein äußerst effizientes Überwachen und Analysieren von Jobs, falls einmal ein Fehler auftreten sollte.

## Graphische Entwicklung

Talend Open Studio's Job Designer bietet sowohl eine graphische als auch eine funktionale Sichtweise des aktuellen Integrations-Prozesses. Der Job Designer enthält eine graphische Darstellung aller verfügbaren Komponenten und Konnektoren – die Komponenten Bibliothek. Integrations-Prozesse werden mittels einfachen Drag and Drop Operationen erstellt. Hierzu werden Komponenten und Konnektoren auf den Arbeitsbereich gezogen, die



Verbindungen graphisch definiert und die Eigenschaften (die meisten Eigenschaften werden aus den Metadaten übernommen) festgelegt.

## Metadatengestütztes Design und Ausführung

Talend Open Studio ist eine von allen Modulen verwendete Metadaten gestützte Lösung, in der alle Informationen in einem Repository gespeichert und verwaltet werden. Dieses zentrale Metadaten Repository verbindet alle Projektinformationen und gewährleistet die Konsistenz über alle Integrations-Prozesse hinweg. Der für die Integrations-Prozesse notwendige Bezug von Quell- und Ziel-System wird durch eine Reihe von Wizards unterstützt. Hierbei werden die Metadaten durch erweiterte Datenbank- und Dateianalysen einfach ins Me-

datanrepository geladen. Das Metadatenrepository beruht auf einem offenen relationalen Modell und gewährleistet die Wartbarkeit durch die Analyse von möglichen Abhängigkeiten und Beeinträchtigungen.

### Zukunftsweisende und vielseitige Konnektivität

Talend Open Studio beinhaltet native und geschäftliche Konnektoren für alle IT-Umgebungen. Dieses breite Spektrum von verschiedensten Treibern ist der Schlüssel für ein erfolgreiches Zusammenarbeiten von Applikationen und Datenbanken. Es erlaubt die Verbindung von unterschiedlichsten und heterogenen Datenstrukturen ohne Geschwindigkeitsverluste. Durch kontinuierliche Erweiterungen wächst der Umfang der Lösung weiter an.

#### Einige Beispiele der enthaltenen Konnektoren:

**Dateien:** CSV, Excel, positional, XML, LDIF, unstrukturiert, etc.

**Datenbanken:** MySQL, MS SQL Server, DB2, Oracle, Ingres, PostgreSQL, Sybase, MS Access, Informix, Firebird, etc.

**Applikationen:** SugarCRM, Salesforce.com, etc.

**Log & error:** info, warning, statistics, filtering, etc.

**Datenqualität:** deduplication, fuzzy logic matching, CRC, etc.

**Transformationen :** filtering, mapping, reference, aggregation, lookup, XSLT, (de)normalizing, etc.

**Misc:** Web Services, FTP, HTTP, system, SSH, data generator, etc.



Eine Übersicht aller verfügbaren Konnektoren ist unter <http://www.talendforge.org/components> abrufbar.

Talend Open Studio beruht auf den von der Industrie standardisierten Sprachen Java, Perl und SQL. Dieses bietet Kunden eine einfache Möglichkeit die bestehenden Komponenten zu verändern oder die Lösung durch Eigenentwicklungen direkt zu ergänzen.

### Real-time debugging

Talend Open Studio beinhaltet leistungsfähige Test-, Debugging- und Tuning-Funktionen welche einen sofortigen Einblick über den gesamten Ablauf zur Laufzeit ermöglichen. Neben einem erweiterten Trace-Mode werden auch direkt Statistiken über die eigentliche Ausführung erzeugt.

Und natürlich ist der von Talend Open Studio erzeugte Programmcode, unabhängig von der Sprache, in der Entwicklungsumgebung immer zugänglich und einsehbar.

### Verteilung und Wartung

Ein spezielles Kontextmanagement (Test, Staging, Produktion, etc.) unterstützt bei der Umsetzung und Entwicklung von Integrations-Prozessen.

USA • Frankreich • Deutschland • China

Die automatische und vollständige Dokumentierung (in XML und HTML) liefert eine immer aktuelle technische Beschreibung.

### Stabile und skalierbare Ausführung

Während viele Integrations-Lösungen auf einem zentralen Integrationsserver beruhen oder die Datenverarbeitung nur innerhalb der Datenbank ausführen können, verteilt Talend Open Studio die Prozesse je nach Verfügbarkeit über den gesamten Rechnerverbund. Der Einsatz von dedizierten Systemen ist hierbei nicht notwendig. Stattdessen werden von Talend Open Studio alle vorhandenen Ressourcen genutzt.

Talend Open Studio ist die einzige Datenintegrationslösung welche sowohl den traditionellen ETL (Extract-Transform-Load) als auch den bekannten ELT (Extract-Load-Transform) Ansatz unterstützt. ELT nutzt die Stärken relationaler Datenbanksysteme indem die Transformationen direkt innerhalb der Datenbank ausgeführt werden. Dadurch können großen Datenmengen in ungeahnter Geschwindigkeit verarbeitet werden. Bei jedem Teilschritt eines Prozesses kann das beste Verfahren für höchste Geschwindigkeit und Skalierbarkeit gewählt werden. Dieses Verfahren, welches speziell für die Verwendung von kostengünstigen Grid-Systemen (Groß oder Klein) entwickelt wurde, erlaubt die Verarbeitung nahe dem Quellsystem und maximiert die Verwendung von Rechenleistung.

*Die Wahl von Talend Open Studio wurde durch zwei technische Aspekte entschieden: Die Lösung ist die beste Wahl für unsere Anforderungen und aus strategischer Sicht gewährleistet das von Talend offerierte Model für Support und Dienstleistungen den besten und langfristigen Schutz in unsere Investition.*  
Toulouse Universität

Talend's Technologie und Geschäftsidee übertrifft die klassischen Modelle der proprietären Anbieter und schafft eine Basis für die sich stetig wechselnden unternehmerischen Anforderungen. Unabhängig von der Unternehmensgröße, dem Wissen oder Budgetbeschränkungen können alle Anforderungen sowohl operativ (Konsolidierung, Datenmigration, etc.) als auch analytisch (ETL, Verdichtung, etc.) umgesetzt werden.

Talend vereinfacht die Datenintegration und macht sie zugänglich für Unternehmen aller Größen und für alle Anforderungen.



[www.talend.com](http://www.talend.com)

[info@talend.com](mailto:info@talend.com)