



IBM InfoSphere Information Server FastTrack: Beschleunigung von Datenintegrations- projekten – ein Muss

Inhalt

- 3 Interaktive Entwicklungsumgebung erstellen**
- 5 Entwicklung von Anwendungen zur Datenintegration optimieren**
- 7 Vorteile auf Unternehmensebene**
- 8 Geschäftsanforderungen bei der Anwendungsimplementierung mit FastTrack überwachen**
 - 9 FastTrack-Projekt starten
 - 12 Profil zu den Quelldaten erstellen
 - 12 Zuordnungsspezifikationen für die Umwandlung und die Bereinigung von Daten erstellen
 - 14 Datenmodell vervollständigen
 - 14 InfoSphere DataStage-Job vervollständigen und implementieren
 - 15 Berichte generieren
- 16 Umfassendere Funktionalität und höhere Produktivität durch die Integration in IBM Module**
 - 17 IBM InfoSphere DataStage und IBM InfoSphere QualityStage
 - 17 IBM InfoSphere Information Analyzer
 - 17 IBM InfoSphere Business Glossary
 - 18 IBM Rational Data Architect
 - 18 IBM Industry Data Models
- 19 Fast-Tracking-Integrationsprojekte mit IBM InfoSphere FastTrack**
- 20 Weitere Informationen**

Damit unternehmensweite Integrationsprojekte allen Geschäftsanforderungen gerecht werden und darüber hinaus innerhalb des vorgesehenen zeitlichen und finanziellen Rahmens durchgeführt werden können, müssen Unternehmensanalysten, Entwickler von Datenmodellen und anderen Lösungen nach Möglichkeiten für eine möglichst effektive Zusammenarbeit finden. Dabei ist die Kommunikation zwischen Einzelpersonen dieser Bereiche besonders wichtig für ein gutes Ergebnis. Häufig wird jedoch ein großer Teil der Entwicklungszeit statt für Programmierung oder Design für das Management und die Umsetzung von Unterschieden hinsichtlich Sprache, Know-how und Arbeitsmethoden sowie für die Zusammenführung von Ergebnissen verschiedener Tools verwendet. Zeitpläne können nicht eingehalten werden, da zu viel Zeit für die Klärung von Geschäftsanforderungen und der Frage verwendet wird, wie die unterstützenden Daten bei der Erstellung der Quelle-Ziel-Zuordnungen verwendet werden. Normalerweise ist die ständige Kommunikation zur Klärung von Bedingungen oder zur Überarbeitung unvollständiger oder unklarer Spezifikationen zeitaufwändiger als erwartet. Letztendlich führen Unterschiede und Neubeginne im Allgemeinen zu Fehlstarts und möglichen Fehleinschätzungen.

Das hohe Maß an manueller Kommunikation während des Collaboration-Prozesses führt nicht nur zu Kommunikations- und Vertrauensproblemen, sondern macht den Prozess darüber hinaus insgesamt ineffizient. In einem typischen Workflow werden vom Entwickler von Datenmodellen Informationsquellen und -ziele entworfen. Diese Strukturen werden anschließend an den Unternehmensanalysten gesendet, der dann manuell die Quelle-Ziel-Zuordnungen definiert, indem er ein Microsoft® Excel-Spreadsheet erstellt, ausdruckt und an den Entwickler sendet. Fehlende Automatisierung und Prüfprotokolle in diesen arbeitsintensiven Prozessen haben zur Folge, dass das Projekt mehr Zeit in Anspruch nimmt, die Zahl der potenziellen Fehlerquellen zunimmt und keine Möglichkeit besteht, Artefakte zu erstellen, die in einem einheitlichen Repository gespeichert und bei nachfolgenden Prozessen von anderen Team-Mitgliedern genutzt werden können.

IBM bietet durch den Einsatz von IBM InfoSphere Information Server FastTrack die Möglichkeit, die Zusammenarbeit im Team zu optimieren, den Grad der Automatisierung zu steigern und sicherzustellen, dass Projekte erfolgreich und pünktlich durchgeführt werden. Durch die Schaffung einer integrierten Umgebung, in der Unternehmensanalysten und Entwickler von Datenmodellen einbezogen werden, beschleunigt FastTrack die gemeinsame Entwicklung von Lösungen, die mehrere Benutzerrollen, Produkte, Länder und Workflows umfassen. Damit wird der Aufwand für mehrere Aufgaben bei der Datenintegration (von der Analyse bis zur Codegenerierung) automatisiert. Hierbei wird die Geschäftsperspektive berücksichtigt, und die Herkunft der Daten sowie dokumentierte Anforderungen bleiben erhalten. FastTrack eignet sich für Unternehmen aller Größen und bietet eine schnellere Wertschöpfung bei strategischen Initiativen, da Geschäftsanforderungen schneller in Datenintegrationsprojekte umgesetzt und damit komplexe geschäftsbezogene Problemstellungen gelöst werden können.

Interaktive Entwicklungsumgebung erstellen

FastTrack ist ein wichtiges Modul der IBM InfoSphere Information Server-Plattform für Datenintegration. Bei der Durchführung der einzelnen Schritten zur Erstellung einer Integrationsanwendung nutzt jedes IBM InfoSphere Information Server-Modul die gemeinsam genutzte Infrastruktur der IBM Information Server-Metadatenservices. Aufgabe des FastTrack-Moduls ist es, den Arbeitsaufwand nachzuvollziehen und zu automatisieren, der bei der Durchführung mehrerer Schritte bei der Datenintegration, wie z. B. Analyse und Codegenerierung, anfällt. Dabei verringert sich der zeitliche Aufwand von der Definition der Geschäftsanforderungen bis zur Implementierung der Lösung.

FastTrack vereinfacht und verbessert den Kommunikationsprozess zwischen dem Unternehmensanalysten, dessen Hauptaugenmerk auf der Geschäftslogik liegt, und dem Entwickler, dessen Hauptaugenmerk auf dem Code liegt. IBM InfoSphere Information Server und FastTrack automatisieren den Informationsfluss zwischen dem Analysten, dessen Aufgabe es ist, Geschäftsanforderungen in eine Reihe von Spezifikationen umzusetzen, die von IT-Mitarbeitern bearbeitet werden können, und dem Entwickler, dessen Aufgabe es ist, diese Anforderungen zu implementieren und eine Anwendung zur Datenintegration zu entwickeln. Dies erfolgt unter anderem durch die Konvertierung eines Excel-Spreadsheets in einen IBM InfoSphere DataStage-Job, der Quelle-Ziel-Zuordnungen grafisch darstellt. Für Entwickler ist es ein erheblich geringerer zeitlicher Aufwand und wesentlich effizienter, einen teilweise abgeschlossenen InfoSphere DataStage-Job, der spezielle Anweisungen für das Abschließen des Jobs enthält, in eine fertig gestellte Anwendung umzuwandeln, als ein Excel-Spreadsheet in das Format derselben Anwendung umzuwandeln.

FastTrack beschleunigt darüber hinaus den Entwicklungsprozess, da es für Unternehmensanalysten den Grad der Automatisierung im Arbeitsablauf erhöht. Der Analyst muss nicht mehr manuell Abhängigkeiten zwischen Daten im Metadatenrepository oder an anderer Stelle nachschlagen. Er kann vielmehr auf FastTrack zurückgreifen, um Abhängigkeiten zwischen Spalten automatisch zu ermitteln und zu vergleichen und den Prozess zur Erstellung der zugehörigen Objektverweise zu automatisieren. Im Rahmen dieses intelligenten Erkennungsprozesses werden Verknüpfungen zwischen Quellen- und Zielstrukturen empfohlen, die auf der Auswertung der Spaltennamen im Hinblick auf exakte, partielle und lexikalische Namensabgleiche basieren. FastTrack kann z. B. unternehmensinterne Standards für Abkürzungen erkennen, wie z. B. „Kunde entspricht Kd.“. Wenn FastTrack die Automatisierung dieses Prozesses übernimmt, kann der Analyst die Zuordnungsspezifikation erheblich schneller und mit weniger Fehlern abschließen. Außerdem ist er in der Lage, die Spezifikation in ein aktives Objekt umzuwandeln, das im Metadatenrepository gespeichert werden soll und von anderen berechtigten Benutzern abgerufen wird, die es zur Verkürzung des Entwicklungszeitraums bei ihren Projekten einsetzen. Das Automatisieren dieses Vorgangs – insbesondere des Erkennungsprozesses – kann in einem ersten Schritt zu einer vollständigeren Spezifikation führen, als durch die manuelle Zuordnung von Datenquellen erreicht werden könnte.

Im Rahmen des automatisierten Prozesses werden wiederverwendbare Artefakte erstellt, die im Metadaten-Service-Repository von IBM InfoSphere Information Server gespeichert werden, und auf die berechnigte Mitarbeiter zugreifen können. Dadurch verbessern sich die Überprüfbarkeit und die Sicherheit bei Integrationsprojekten. Bisher waren Projektspezifikationen auf zahlreiche verschiedene Stellen in unterschiedlichen Formaten verteilt. Mit Hilfe von FastTrack kann das Prüfprotokoll auf einfache Weise eingehalten werden, und Fragen zu langfristigen Entscheidungen können schnell und zuverlässig beantwortet werden.

Entwicklung von Anwendungen zur Datenintegration optimieren

Der zentrale Faktor bei der Optimierung von Aufwänden in der Entwicklung ist das Konzept, mit dem die Umwandlung von Geschäftsanforderungen in fertig gestellte Anwendungen beschleunigt und automatisiert werden soll. FastTrack wurde vor diesem Hintergrund entwickelt und weist spezielle Features im Hinblick auf die Reduzierung manueller Aufwände während der Entwicklungszyklen auf ein Minimum auf. Dennoch wird weiterhin ein Höchstmaß an Auswirkungen auf nachgelagerte Entwicklungsaktivitäten erreicht. Zu diesen Features gehören:

- **Automatisierte Generierung von IBM InfoSphere DataStage-Jobs** – Automatisiert die gleichzeitige Generierung mehrerer Jobs und bietet die Möglichkeit, mehrere Spezifikationen zur Erstellung eines einzelnen Jobs miteinander zu kombinieren; leitet automatisch einen ETL-Fluss (Extract/Transform/Load) von Geschäftsanforderungen ab.
- **Anpassbare Spreadsheet-Sicht** – Bietet den Benutzern die Möglichkeit, Geschäftsregeln und/oder Transformationslogiken als Anmerkungen in Spaltenzuordnungen aufzunehmen und detaillierte Informationen zu Eigenschaften anzuzeigen; dies umfasst auch eine erweiterbare Sicht der Artefakte im Metadaten-Service-Repository von IBM InfoSphere Information Server.
- **Automatisierte Erkennung von Abhängigkeiten** – Umfasst anpassbare Erkennungsalgorithmen zur Ermittlung exakter, partieller und lexikalischer Übereinstimmungen bei sich entsprechenden Spaltennamen in Quellen- und Zielstrukturen, die in einem einfachen und benutzerfreundlichen Format dargestellt werden.

- **Begriffe in IBM InfoSphere Business Glossary aufnehmen und mit Ressourcen verknüpfen** – Bietet Analysten die Möglichkeit, neue Geschäftsbedingungen festzulegen und deren Beziehung zu den entsprechenden physischen Spalten als Teil des Zuordnungsprozesses zu dokumentieren; diese Informationen werden anschließend im gemeinsamen Repository für IBM InfoSphere Information Server Metadaten-Services veröffentlicht, um allen Entwicklungsteams eine vollständige Sicht auf die Geschäftsprozesse zu bieten.
- **Einfacher Import von Daten aus Excel-Spreadsheets und .csv-Formaten** – Ermöglicht den Benutzern den Import von Daten aus vorhandenen Spreadsheets, um neue Spezifikationen anzulegen; außerdem können Daten in Excel-Spreadsheets und .csv-Formate exportiert werden.
- **Flexibel einsetzbare Berichtsfunktionen** – Bietet Benutzern die Möglichkeit, Details zu Zuordnungen auf Spaltenebene und auf Tabellenebene sowie zu Suchfunktionen anzuzeigen; verwaltet und überwacht Metadaten während des gesamten Integrationszyklus; bietet Funktionen zum Erstellen von Berichten im PDF- und HTML-Format auf der Grundlage der Spezifikationen, die anschließend Entwicklungsteams und dem Management zur Verfügung gestellt werden können.
- **Zuordnungen auf Tabellenebene** – Ermöglicht den Benutzern die Definition von Verknüpfungen, Filtern und Vorgangsanweisungen und die Nutzung von Verknüpfungen, die während der Profilerstellung mit IBM InfoSphere Information Analyzer ermittelt wurden.
- **Unterstützung von Suchfunktionen** – Bietet die Möglichkeit, Suchen zu definieren und die „Suchphase“ in dem generierten Integrationsjob zu generieren; diese Suchdefinitionen können durchsucht und wiederverwendet werden.
- **DDL (Data Definition Language) generieren** – Option zum Generieren eines Zieldatenmodells, auf das von Entwicklern von Datenmodellen zurückgegriffen werden kann, falls derzeit keines für die Zielanwendung vorhanden ist.

Vorteile auf Unternehmensebene

Durch die Schaffung einer höchst effizienten, automatisierten Umgebung für die Entwicklung von Integrationsanwendungen und den zugehörigen Designspezifikationen bietet InfoSphere FastTrack Unternehmen eine Reihe von Vorteilen. Hierzu gehören:

- **Bessere Onlinezusammenarbeit** zwischen Unternehmensanalysten, Entwicklern und Datenadministratoren, da der Einsatz von Artefakten, die für jeden Aufgabenbereich generiert wurden, vereinfacht wird. Außerdem verringern sich manuelle Aufwände, falsche Interpretationen und Unklarheiten.
- **Schnellere Wertschöpfung** bei Geschäftsprojekten durch die Automatisierung der zugrundeliegenden Datenintegrationsprozesse.
- **Höhere Erfolgsquoten bei strategischen Initiativen**, da ein einheitliches Verständnis der Geschäftsanforderungen und eine effiziente Integration in technische Lösungen sichergestellt sind.
- **Weniger Auswirkungen bei Veränderungen** aufgrund der End-to-End-Funktionen zur Herkunft der Daten und zum Management von Metadaten auf der IBM InfoSphere Information Server-Plattform.
- **Bessere Daten-Governance und Prüffunktionen** durch die Bereitstellung eines zentralen, verwalteten Standorts für Projektspezifikationen und -anforderungen.
- **Bessere Überprüfbarkeit** und Zugriff auf die Herkunftsstruktur zwischen Jobs und Spezifikationen.

Geschäftsanforderungen bei der Anwendungsimplementierung mit FastTrack überwachen

Der in der folgenden Abbildung dargestellte allgemeine Arbeitsablauf veranschaulicht, wie ein Unternehmen FastTrack als Teil der IBM Information Server-Plattform einsetzen kann, um die Integration der Anwendungsentwicklung zu beschleunigen und zu automatisieren (siehe Abbildung 1). Die Sicherheitsrollen werden hierbei während des gesamten Entwicklungsprozesses beibehalten. Das heißt, die Benutzer benötigen speziell zugeordnete Berechtigungen, um Einträge in InfoSphere Business Glossary anzeigen oder ändern zu können, oder um Profile zu Datenquellen mit InfoSphere Information Analyzer erstellen zu können.

Abbildung 1: FastTrack kommt bei der Umwandlung von Geschäftsanforderungen in ausführbaren Code für die Anwendungsimplementierung strategische Bedeutung zu.

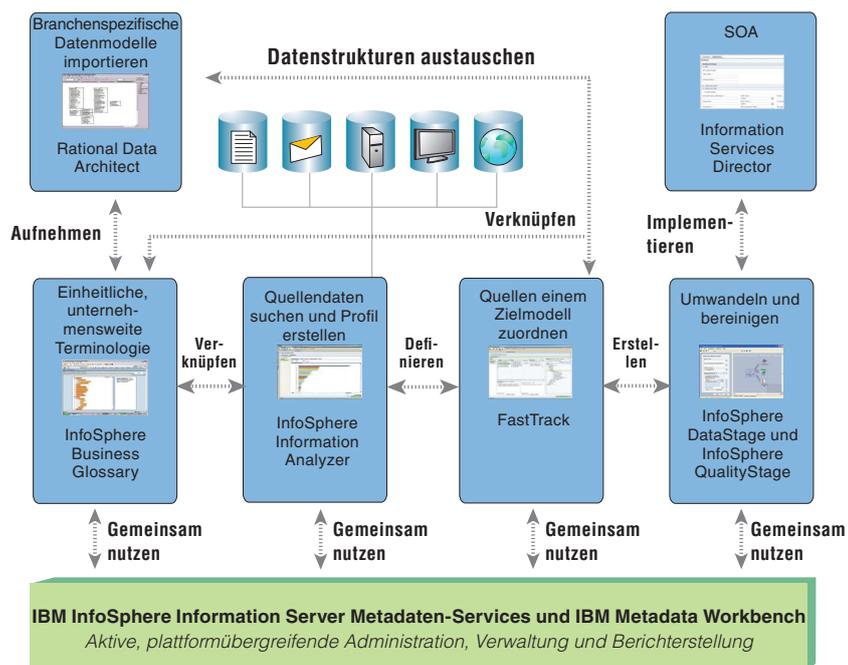


Abbildung 2: Benutzer können Daten aus vorhandenen Excel-Spreadsheets oder .csv-Formaten importieren.

FastTrack-Projekt starten

Benutzer können ein FastTrack-Projekt durch Importieren eines vorhandenen Excel-Spreadsheets (siehe Abbildung 2) oder von Grund auf starten und die Anforderungen direkt in die FastTrack-Schnittstelle

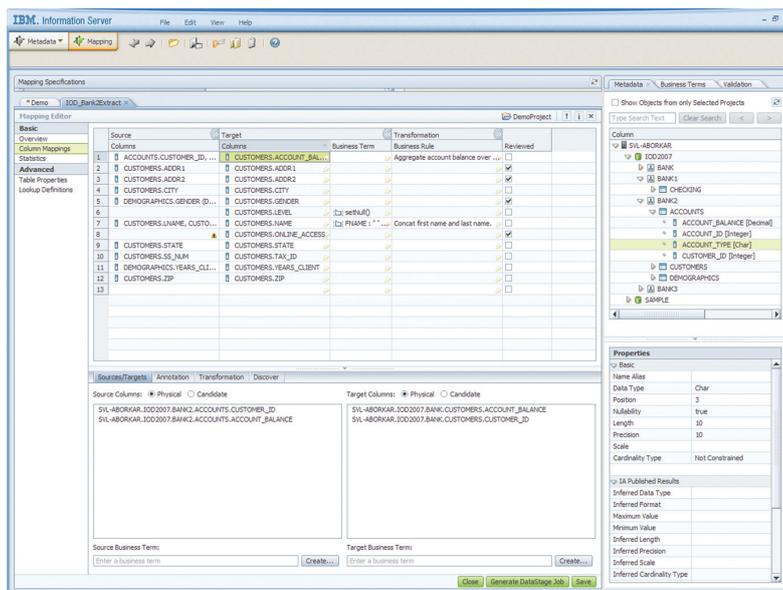
The screenshot displays a Microsoft Excel spreadsheet titled '100_Bank2Extract.xls'. The spreadsheet contains a table with the following columns: Source Columns, Target Columns, Transformation Rule, Transformation Function, and Annotation. The rows define mappings for customer data, including fields like FNAME, LNAME, ADDR1, ADDR2, CITY, STATE, ZIP, ACCOUNT_BALANCE, TAX_ID, YEARS_CLIENT, GENDER, ONLINE_ACCESS, and LEVEL.

Below the spreadsheet is a 'Demo' dialog box for 'Import from CSV'. It includes options for 'Select the comma separated values import options.', 'Start import from row:' (set to 1), and 'Delimiter:' (set to comma). There are checkboxes for 'Omit Hostname from Multipart Data Source Names', 'Case Sensitive Matching for Data Source Names', and 'Use First Row as Column Headers'. A table below these options shows a preview of the source and target columns with their respective transformation rules and functions. At the bottom, there are dropdown menus for selecting source and target columns, business terms, and transformation rules, along with checkboxes for 'Join' and 'Filter'.

Source Columns	Target Columns	Transformation Rule	Transformation Function	Annotation
iod BANK2 CUSTOMERS.FNAME,iod BANK2 CUSTOMERS.LNAME	iod BANK CUSTOMERS.NAME	Concat first name and last name.	FNAME & " " & LNAME	
iod BANK2 CUSTOMERS.ADDR1	iod BANK CUSTOMERS.ADDR1			
iod BANK2 CUSTOMERS.ADDR2	iod BANK CUSTOMERS.ADDR2			
iod BANK2 CUSTOMERS.CITY	iod BANK CUSTOMERS.CITY			
iod BANK2 CUSTOMERS.STATE	iod BANK CUSTOMERS.STATE			
iod BANK2 CUSTOMERS.ZIP	iod BANK CUSTOMERS.ZIP			
iod BANK2 ACCOUNTS.ACCOUNT_BALANCE,iod BANK2 ACCOUNTS.CUSTOMER_ID	iod BANK CUSTOMERS.ACCOUNT_BALANCE,iod BANK CUSTOMERS.CUSTOMER_ID	Aggregate account balance over savings and checking accounts by customer id to compute the total account balance for each customer.		
iod BANK2 DEMOGRAPHICS.YEARS_CLIENT	iod BANK CUSTOMERS.YEARS_CLIENT			
iod BANK2 DEMOGRAPHICS.GENDER	iod BANK CUSTOMERS.GENDER			
	iod BANK CUSTOMERS.ONLINE_ACCESS			
	iod BANK CUSTOMERS.LEVEL		setNull()	Will be computer after standardization.

eingeben (siehe Abbildung 3). Wenn sie ein Projekt von Grund auf starten, kann der Analyst den Integrationsprozess beschleunigen, indem er die IBM Industry Data Models nutzt, um das Ziel-Data-Warehouse oder die Zielanwendung zu definieren. Außerdem können die zuvor im Rahmen der gemeinsam benutzten IBM InfoSphere Information Server-Metadatenervices gesammelten und gespeicherten Metadaten dazu verwendet werden, die Erstellung von Quelle-Ziel-Zuordnungen weiter zu beschleunigen. Mit diesen Zuordnungen wird definiert, wie Daten aus einer Quelle extrahiert, auf der Basis von Geschäftsregeln umgewandelt und in das Ziel-Data-Warehouse oder die Zielanwendung geladen werden. Mit Hilfe der Funktionalität für den Import von Metadaten in FastTrack besteht darüber hinaus die Möglichkeit, neue Metadaten von Quellsystemen zu importieren, die bisher nicht in IBM InfoSphere Information Server definiert wurden.

Abbildung 3: FastTrack-Schnittstelle.



Zu den zentralen Anforderungen bei einem Datenintegrationsprojekt gehören die Definition und die Dokumentation unternehmensweiter Standards, die für Definitionen, Business-Transformation-Regeln und zugrundeliegenden physischen Datenstrukturen angewendet werden sollen.

Unternehmensanalysten müssen häufig einheitliche Definitionen festlegen oder nutzen, um eine einheitliche Terminologie in mehreren Entwicklungsteams zu gewährleisten. Die Analysten können bei der Festlegung dieser Definitionen von Grund auf beginnen. Falls jedoch unternehmensweit verwendete Geschäftsbegriffe bereits extern festgelegt wurden oder auf die IBM Industry Data Models zurückgegriffen wurde, um die Durchführung von Aufgaben bei der Projektentwicklung zu beschleunigen, können diese Geschäftsinformationen auch in InfoSphere Business Glossary importiert und direkt im FastTrack-Framework genutzt werden. Mit FastTrack werden nicht nur einheitliche Definitionen festgelegt, sondern Analysten haben darüber hinaus die Möglichkeit, vorhandene und neu festgelegte Geschäftsbegriffe in vollem Umfang zu nutzen. Hierbei werden diese Geschäftsbegriffe mit den zugehörigen physischen Strukturen verknüpft, um deren umfassende Abhängigkeiten zu veranschaulichen. Die gesamte Herkunftsstruktur von den geschäftsspezifischen bis zu den technischen Metadaten kann anschließend in InfoSphere Business Glossary veröffentlicht werden, auf das Benutzer mit den entsprechenden Berechtigungen zugreifen können.

Während des Prozesses der Zuordnung von Spezifikationen müssen Business-Transformation-Regeln eindeutig dokumentiert werden, um dem Entwickler die Geschäftsziele in einer einheitlichen Sprache zu verdeutlichen. Die Analysten können bei Bedarf manuelle Zuordnungen hinzufügen, um diese kritische Anforderung zu erfüllen. Sie können außerdem spezielle Funktionen nutzen (z. B. Erkennungsfunktionen), mit denen das Metadaten-Repository nach Übereinstimmungen zwischen definierten Quellen- und Zielsystemen auf Spaltenebene durchsucht wird. Diese Informationen werden gesammelt, protokolliert und dem Entwickler im Rahmen des Prozesses der Zuordnung von Spezifikationen dargestellt.

Der Unternehmensanalyst ist außerdem dafür verantwortlich, dass während der Erstellung dieser detaillierten Zuordnungsspezifikationen die Namenskonventionen eingehalten werden. Bei der Bezugnahme auf einen Kunden empfiehlt es sich zum Beispiel, unternehmensweit die Kurzform „Kd“ zu verwenden. Für vorhandene oder neu erstellte Datenmodelle können Namensstandards physischer Datenstrukturen angewendet werden, um eine einheitliche Benennung für alle unternehmensweit eingesetzten Anwendungen zu gewährleisten.

Profil zu den Quelldaten erstellen

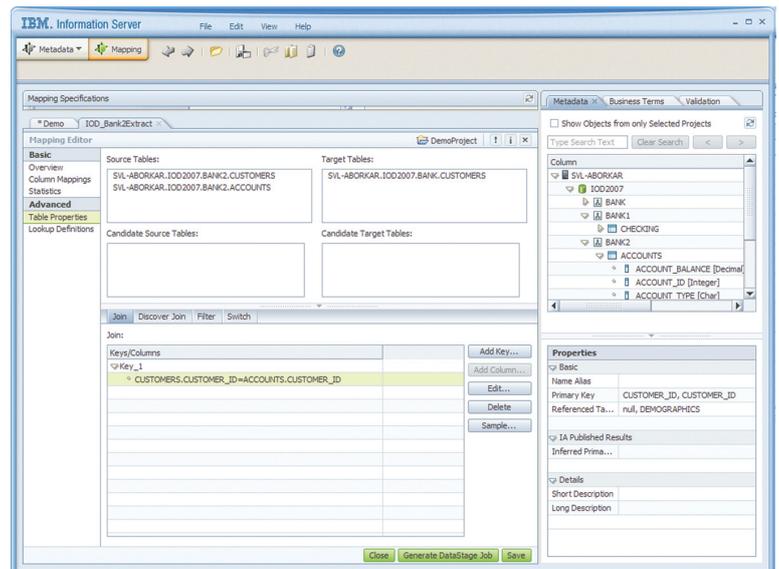
Datenanalysten und SMEs (Subject Matter Experts) erstellen Profile zu Daten aus mehreren Quellen, um ein Verständnis für die Arten von Transformationsregeln zu erhalten, die während des Migrationsprozesses angewendet werden müssen. Auf der Grundlage von Inhalten aus InfoSphere Information Analyzer und InfoSphere Business Glossary dokumentieren die Analysten alle Ergebnisse in Notes, die anschließend dem FastTrack-Benutzer zur Verfügung gestellt und in die Eigenschaftsanzeige aufgenommen werden, damit auf einfache Weise darauf zugegriffen werden kann. Dieser Prozess läuft häufig schrittweise ab. Die Möglichkeit, diese Metadaten aktiv über IBM InfoSphere Information Server-Metadatenservices zur Verfügung zu stellen, ist daher ein kritischer Faktor für die Optimierung dieses Prozesses und die Gewährleistung, dass keine Details übersehen werden.

Zuordnungsspezifikationen für die Umwandlung und die Bereinigung von Daten erstellen

Bei der Arbeit mit FastTrack steht dem Unternehmensanalysten eine Spreadsheet-ähnliche Schnittstelle zur Eingabe von Quelle-Ziel-Zuordnungsspezifikationen zur Verfügung. Mit diesen Spezifikationen kann z. B. beschrieben werden, wie Informationen aus Quellen extrahiert, Daten gefiltert, Quellenspalten Geschäftsbedingungen zugeordnet und diese Begriffe dazu verwendet werden, um ein neues oder aktualisiertes Zieldatenmodell sowie die zugehörigen Definitionen zur Tabellenerstellung zu generieren. Die Zuordnungen können recht komplex sein und auf mehreren Ebenen erstellt werden.

Mit Hilfe von Zuordnungen auf Tabellenebene (siehe Abbildung 4) haben Analysten die Möglichkeit, Filterkriterien für Datengruppen festzulegen und Suchinformationen zu definieren, die aus anderen Datenentitäten extrahiert werden sollen. Die Analysten können so neue Wege zur Verknüpfung von Tabellen in FastTrack erarbeiten oder auf die Ergebnisse der Profilerstellung aus InfoSphere Information Analyzer zurückgreifen, um zu definieren, wie mehrere Tabellen miteinander verknüpft werden müssen. Mit Zuordnungen auf Spaltenebene werden spezielle Transformationsregeln definiert, wie z. B. das Festlegen von Standardwerten, das Prüfen gültiger Werte, die Durchführung mathematischer Funktionen und von Zeichenfolgeverarbeitungen sowie viele andere Möglichkeiten. Wenn der Analyst nicht sicher ist, wie eine bestimmte Geschäftsregel angewendet werden soll, kann er detaillierte Notizen und Anmerkungen für eine Spalte anlegen und einsetzen, die dem Entwickler über die Zuordnungsspezifikation und die generierten InfoSphere DataStage-Jobs automatisch als Teil der Geschäftsanforderungen zur Verfügung gestellt werden. Während des Zuordnungsprozesses kann der Analyst außerdem die Zuordnungen von Geschäftsbedingungen auf Spaltenebene definieren. Andere Benutzer können sich so mit dem gesamten Spektrum an Informationen (vom Standpunkt des Unternehmens bis zur technischen Implementierung) vertraut machen.

Abbildung 4: Zuordnungen auf Tabellenebene.



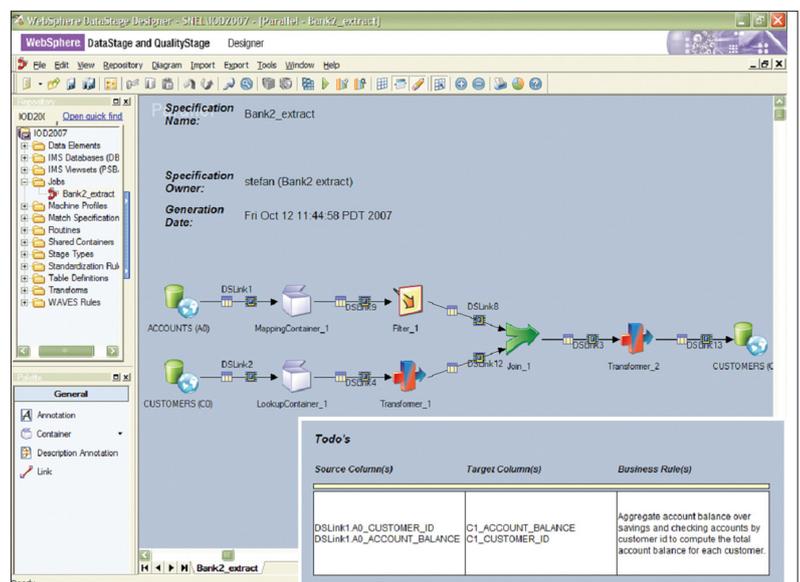
Datenmodell vervollständigen

Sobald die Zuordnungsspezifikation vollständig ist, wird die aktualisierte Zielmodellstruktur in die Spezifikationsdetails aufgenommen. An diesem Punkt kann der Analyst ein neues oder aktualisiertes DDL-Script (Data Definition Language) generieren. Dieses Script wird zur Prüfung, Erweiterung und Implementierung in die Produktionsanwendung an den Datenadministrator und die Entwickler der Datenmodelle weitergeleitet. Während dieses Schritts können alle vorhandenen Namensstandards und -konventionen automatisch angewendet werden. Der Datenadministrator und die Entwickler des Datenmodells müssen daher möglicherweise vor der Implementierung nur wenige Änderungen vornehmen.

InfoSphere DataStage-Job vervollständigen und implementieren

Der Analyst kann jetzt InfoSphere DataStage-Jobs generieren und an den Entwickler zur Prüfung und zur Vervollständigung im Hinblick auf die Produktimplementierung übergeben. Mit nur einem Mausklick wandelt der Analyst die Geschäftslogik aus einer oder mehreren FastTrack-Spezifikationen in einen parallelen InfoSphere DataStage ETL-Job um, in dem der Datenfluss und die zugehörigen Regeln grafisch dargestellt werden (siehe Abbildung 5).

Abbildung 5: In FastTrack generierter InfoSphere DataStage-Job.



Der generierte Job enthält alle vom Analysten eingegebenen Transformationsregeln. Die erfasste Geschäftslogik verläuft daher ohne manuellen Eingriff automatisch bis zur Implementierung. Diese Jobs können häufig umgehend kompiliert und unverändert ausgeführt werden. Falls vor der Implementierung einer Produktionsanwendung eine zusätzliche Logik oder eine weitere Optimierung erforderlich ist, hat der ETL-Entwickler die Möglichkeit, den Job zu bearbeiten und alle relevanten Notes vom Analysten zu lesen, die direkt als Anmerkungen in der grafischen Jobanzeige angezeigt werden.

Berichte generieren

FastTrack-Benutzer können eine Vielzahl von Berichten generieren, wie z. B. einen End-to-End-Bericht zur Datenumwandlung mit allen definierten Spezifikationen, die von anderen Mitgliedern des Projektteams angezeigt werden können. Die Berichte werden über das IBM InfoSphere Information Server Core Services Reporting Framework generiert. Hierbei handelt es sich um eine webbasierte Konsole, auf der alle Plattformberichte veröffentlicht werden. In Berichte können Informationen zu Tabellen, Zuordnungen auf Tabellen- und Spaltenebene sowie Anmerkungen aufgenommen werden, die anschließend als Dokumente im PDF- oder HTML-Format veröffentlicht werden können (siehe Abbildung 6).

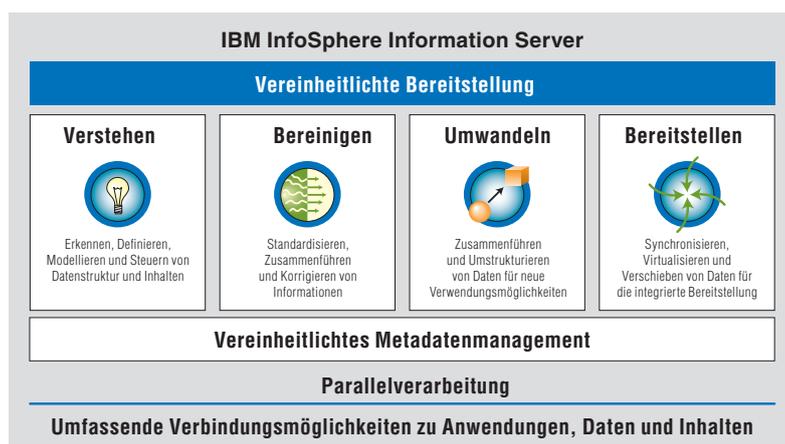
Abbildung 6: Der End-to-End-Bericht enthält Details zu Zuordnungen und Suchen auf Spalten- und Tabellenebene.

Mapping Specification	
Implementation Details	
Project :	IOD
Report Name :	Bank2Extract_details
Report Generated :	2007-10-11 07:39:26
User :	admin admin
Description :	mapping specification, table mapping, column mapping, and lookup details
Specification Name :	IOD_BankExtract2_Finished
Owner :	
Status :	In Progress
Mapping Specification Description :	
Details	
Source Table :	iod.iod.BANK2.ACCOUNTS iod.iod.BANK2.CUSTOMERS iod.iod.BANK2.DEMOGRAPHICS
Target Table :	iod.iod.BANK.CUSTOMERS
Filter :	ACCOUNT_BALANCE <>-999999
Join :	iod.iod.BANK2.CUSTOMERS.CUSTOMER_ID = iod.iod.BANK2.CUSTOMERS.CUSTOMER_ID
Case :	
Reviewed :	false
Documentation :	
Details	
Source Column :	iod.iod.BANK2.ACCOUNTS.ACCOUNT_BALANCE iod.iod.BANK2.ACCOUNTS.CUSTOMER_ID
Source Term :	
Target Column :	iod.iod.BANK.CUSTOMERS.ACCOUNT_BALANCE iod.iod.BANK.CUSTOMERS.CUSTOMER_ID
Target Term :	
Transformation :	
Business Rule :	Aggregate account balance over savings and checking accounts by customer id to compute the total account balance for each customer.
Reviewed :	false
Documentation :	
All Rights Reserved. IBM is a trademark of IBM Corp. and IBM Information Server is a trademark or registered trademark in the United States, other countries or both.	

Umfassendere Funktionalität und höhere Produktivität durch die Integration in IBM Module

IBM InfoSphere Information Server ist eine integrierte Plattform, die auf dem gemeinsamen Repository der IBM InfoSphere Information Server-Metadatenservices basiert. Die Software enthält spezielle Module für jeden Prozessschritt zur Entwicklung einer Datenintegrationsanwendung (siehe Abbildung 7). FastTrack bietet im Entwicklungsprozess den größten Nutzen, wenn es in Kombination mit den übrigen Komponenten der IBM InfoSphere Information Server-Plattform zur Datenintegration verwendet wird. Zur Integration von FastTrack in den Entwicklungsprozess sollten Unternehmen IBM InfoSphere Information Server als Standard für den Datenintegrationsprozess nutzen. FastTrack generiert automatisch IBM InfoSphere DataStage-Jobs, die darauf ausgerichtet sind, die komplexen Anforderungen bei ETL-Verarbeitungen zu erfüllen. Für viele Unternehmen sind diese Anforderungen als grundlegende Komponente ihrer Integrationsarchitektur unabdingbar. Andere Module können wahlweise installiert werden. Durch das Hinzufügen der in diesem Dokument beschriebenen Module erhöht sich allerdings der Grad an Automatisierung, Effizienz und Funktionalität im gesamten Entwicklungsprozess.

Abbildung 7: IBM InfoSphere Information Server – eine Plattform zur Datenintegration, über die zuverlässige Informationen zu wichtigen Geschäftsinitiativen bereitgestellt werden.



IBM InfoSphere DataStage und IBM InfoSphere QualityStage.

Um die Vorteile von FastTrack in vollem Umfang zu nutzen, können Benutzer ihre Jobspezifikationen direkt in ETL-Jobs konvertieren, die über InfoSphere DataStage implementiert werden. Dies ist ein Modul von IBM InfoSphere Information Server, das große Datenmengen in ein geeignetes gängiges Format für den Anwendungszugriff visuell umwandelt und zusammenstellt. Es zeichnet sich durch codefreie Top-Down-Darstellungen von Datenflüssen mit Hunderten integrierter Umwandlungsfunktionen aus, wie z. B. erweiterte Verfahren zur Datenbereinigung auf der Grundlage von IBM InfoSphere QualityStage. InfoSphere DataStage basiert nativ auf einer exklusiv parallelen Transformation Engine und unterstützt Verarbeitungsanforderungen im Batchbetrieb, in Echtzeit und auf der Grundlage von SOA (Service Oriented Architecture).

IBM InfoSphere Information Analyzer. Diese Software ermöglicht ein grundlegendes Verständnis des Inhalts, der Qualität und der Struktur von Quelldaten. InfoSphere Information Analyzer bietet Funktionen für die Spaltenanalyse auf der Basis tatsächlicher Datenwerte (keine Metadatenwerte), ermittelt die wirklichen physischen Merkmale der Daten (z. B. Datentyp, Genauigkeit, Umfang, Ungültigkeit und andere), berechnet die Häufigkeitsverteilung, identifiziert unterschiedliche Werte und erstellt eine Datei mit Beispieldaten. Primärschlüssel, Fremdschlüssel und redundante Domänenanalyse sind ebenfalls Bestandteil des Prozesses zur Profilerstellung. InfoSphere Information Analyzer ersetzt den zeitaufwändigen und fehleranfälligen Prozess der manuellen Analyse von Datenprofilen.

IBM InfoSphere Business Glossary. Bei dieser Software handelt es sich um ein webbasiertes Tool, mit dem Unternehmensanalysten und SMEs (Subject Matter Experts) ein einheitliches, unternehmensweites Terminologie- und Klassifizierungssystem erstellen, verwalten und verteilen können. InfoSphere Business Glossary ist aktiv mit den IBM InfoSphere Information Server-Metadatenservices verknüpft und bietet Datenadministratoren die Möglichkeit, Geschäftsbedingungen mit technischen Artefakten zu verknüpfen, die wiederum mit anderen Modulen von IBM InfoSphere Information Server gemeinsam benutzt werden. Dadurch ergeben sich einheitliche semantische Tags, die von Entwicklern von Datenmodellen, Datenanalysten, Unternehmensanalysten und Endbenutzern verwendet werden.

IBM Rational Data Architect. Rational Data Architect (RDA) ist ein unternehmensweit eingesetztes Tool zur Entwicklung und Integration von Datenmodellen mit Funktionen für die Entwicklung von logischen und physischen Modellen sowie von Glossarmodellen. RDA wurde speziell konzipiert, um Datenadministratoren dabei zu unterstützen, ein Verständnis für IT-Ressourcen zu bekommen, ein Modell der zugehörigen Abhängigkeiten zu erarbeiten, miteinander verbundene Datenbanken zu entwickeln und Datenbankprojekte zu optimieren. Physische und glossarspezifische Informationen können mit den IBM Information Server-Metadatenservices zur Verwendung in Datenintegrationsprojekten ausgetauscht werden.

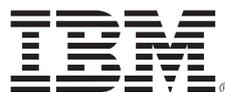
IBM Industry Data Models. Die IBM Industry Data Models können sowohl von Business- als auch von IT-Seite verwendet werden, um wichtige strategische Geschäftsinitiativen schneller und zuverlässiger zu implementieren. Grundlage dieser Software sind ein umfassendes branchenspezifisches Know-how mit mehr als 400 Kunden und über zehn Jahre Erfahrung in der Entwicklung. Die IBM Industry Data Models können daher in einer Vielzahl von Branchen eingesetzt werden. Hierzu gehören:

- *IBM Banking Data Warehouse (BDW)*
- *IBM Insurance Information Warehouse (IIW)*
- *IBM Retail Data Warehouse (RDW)*
- *IBM Telecommunications Data Warehouse (TDW)*
- *IBM Health Plan Data Warehouse (HPDW)*
- *IBM Financial Markets Data Warehouse (FMDW)*

Zu den Datenmodellen gehört ein Glossar mit Begriffen und Konzepten, das leicht verständlich ist und von Business- und IT-Benutzern über IBM InfoSphere Business Glossary gemeinsam benutzt werden kann. Dadurch können Abgrenzungen eines Projekts beschleunigt, geeignete Berichte erstellt, Datenanforderungen festgelegt und Datenquellen definiert werden.

**Fast-Tracking-Integrationsprojekte mit IBM InfoSphere Information Server
FastTrack**

Im Hinblick auf die komplexen Integrationsprojekte, die im heutigen Geschäftsumfeld von Unternehmen aller Größen durchgeführt werden, verkürzt FastTrack den Zeitraum von der Zusammenstellung der Anforderungen bis zur Implementierung. Hierbei wird die Zusammenarbeit zwischen den Benutzern an mehreren Standorten und mit unterschiedlichem Know-how vereinfacht und verbessert, und Aufgaben, die bisher manuell durchgeführt werden mussten, werden automatisiert. Als Teil der Integrationsplattform, die Benutzern von IBM InfoSphere Information Server zur Verfügung steht, bietet FastTrack Unternehmen die Möglichkeit, innerhalb kurzer Zeit integrierte Anwendungen zu entwickeln, die kritische Geschäftsanforderungen erfüllen und Wettbewerbsvorteile ermöglichen.



Weitere Informationen

Weitere Informationen über IBM InfoSphere Information Server und FastTrack erhalten Sie über den zuständigen IBM Ansprechpartner oder auf folgender Website:

ibm.com/software/data/ips

IBM Deutschland GmbH
70548 Stuttgart
ibm.com/de

IBM Österreich
Obere Donaustraße 95
1020 Wien
ibm.com/at

IBM Schweiz
Vulkanstrasse 106
8010 Zürich
ibm.com/ch

Die IBM Homepage finden Sie unter:

ibm.com

IBM, das IBM Logo und ibm.com sind eingetragene Marken der IBM Corporation.

DataStage, QualityStage, Rational und InfoSphere sind Marken der IBM Corporation in den USA und/oder anderen Ländern.

Microsoft ist eine Marke der Microsoft Corporation in den USA und/oder anderen Ländern.

Weitere Unternehmens-, Produkt- oder Servicenamen können Marken anderer Hersteller sein.

Vertragsbedingungen und Preise erhalten Sie bei den IBM Geschäftsstellen und/oder den IBM Business Partnern. Die Produktinformationen geben den derzeitigen Stand wieder. Gegenstand und Umfang der Leistungen bestimmen sich ausschließlich nach den jeweiligen Verträgen. Angebote können jederzeit ohne vorherige Ankündigung geändert, erweitert oder zurückgezogen werden.

Die oben genannten Erklärungen bezüglich der Produktstrategien und Absichtserklärungen von IBM stellen die gegenwärtige Absicht von IBM dar, unterliegen Änderungen oder können zurückgenommen werden und stellen nur die Ziele von IBM dar.

Der Inhalt dieser Dokumentation dient nur zu Informationszwecken. Obwohl die in dieser Dokumentation enthaltenen Informationen auf ihre Vollständigkeit und Genauigkeit hin überprüft wurden, werden sie auf der Grundlage des gegenwärtigen Zustands (auf „as-is“-Basis) ohne jegliche Gewährleistung zur Verfügung gestellt. Darüber hinaus basieren diese Informationen auf den derzeitigen Produktplänen und -strategien von IBM, die jederzeit ohne vorherige Ankündigung geändert werden können. IBM übernimmt keine Haftung für irgendwelche Schäden, die aus der Nutzung dieser oder einer anderen Dokumentation entstehen oder damit in Zusammenhang stehen. Aus dem Inhalt dieser Dokumentation können kein Gewährleistungsanspruch oder andere Anforderungen an IBM (oder seine Lieferanten oder Lizenzgeber) abgeleitet werden, noch kann der Inhalt eine Änderung der Bedingungen der geltenden Lizenzvereinbarung, der die Nutzung der IBM Software unterliegt, bewirken.

© Copyright IBM Corporation 2008
Alle Rechte vorbehalten.

TAKE BACK CONTROL WITH **Information Management**

IMW14001-DEDE-00