

Stammdatenmanagement

Dr. Christine Legner / Dr. Boris Otto, St. Gallen

Die Grunddaten eines Unternehmens, z.B. Kunden, Artikel, Lieferanten oder Mitarbeiter, sind für die Abwicklung integrierter Geschäftsprozesse von zentraler Bedeutung. Informationssysteme repräsentieren diese in Form von Stammdaten, die in der Regel über einen längeren Zeitraum unverändert bleiben. Im Zuge der Verteilung von Geschäftsprozessen innerhalb eines Unternehmens oder über Unternehmensgrenzen hinweg steigen die Anforderungen an die Sicherstellung der Stammdatenqualität, die Pflege von Stammdaten sowie deren Abgleich zwischen verschiedenen Informationssystemen. Der vorliegende Beitrag fasst Grundlagen des Stammdatenmanagements zusammen und geht insbesondere auf die Stammdatenqualität, -architektur und -organisation ein.

1. Definition

<RN>Stammdaten</RN>

Als Stammdaten bezeichnet man in der betrieblichen Datenverarbeitung die wesentlichen Grunddaten eines Unternehmens, die sich durch operative Prozesse überhaupt nicht oder nur selten verändern, während die durch die laufende Geschäftsabwicklung anfallenden Daten als Dispositionsdaten bezeichnet werden. Zu den Stammdaten zählen beispielsweise Kunden-, Material-, Mitarbeiter- und Lieferantendaten. Ein Materialstammsatz speichert beispielsweise sämtliche Informationen zu den Artikeln, Teilen und Dienstleistungen, die ein Unternehmen beschafft, fertigt und lagert. Er wird von den verschiedenen funktionalen Bereichen (z.B. Einkauf, Verkauf, Logistik, Produktion oder Buchhaltung) genutzt. Diese benötigen neben den allgemeinen Grunddaten, wie z.B. einer Materialnummer zur eindeutigen Identifikation und einer Bezeichnung, in der Regel noch funktionspezifische Informationen, die in Sichten gruppiert werden: So müssen dem Einkauf insbesondere Preise und Konditionen zur Verfügung stehen, während für die Fertigungsplanung (Disposition) Sicherheitsbestände und Losgrößen relevant sind und die Buchhaltung vor allem an Kontierungsinformationen interessiert ist. Stammdaten können von der Organisationsstruktur abhängig sein bzw. in mehreren Sprachen gepflegt werden (s. Abbildung 1). Durch die Integration zum Beispiel aller materialspezifischen Informationen in einem einzigen Stammsatz entfällt dafür die redundante Datenhaltung.

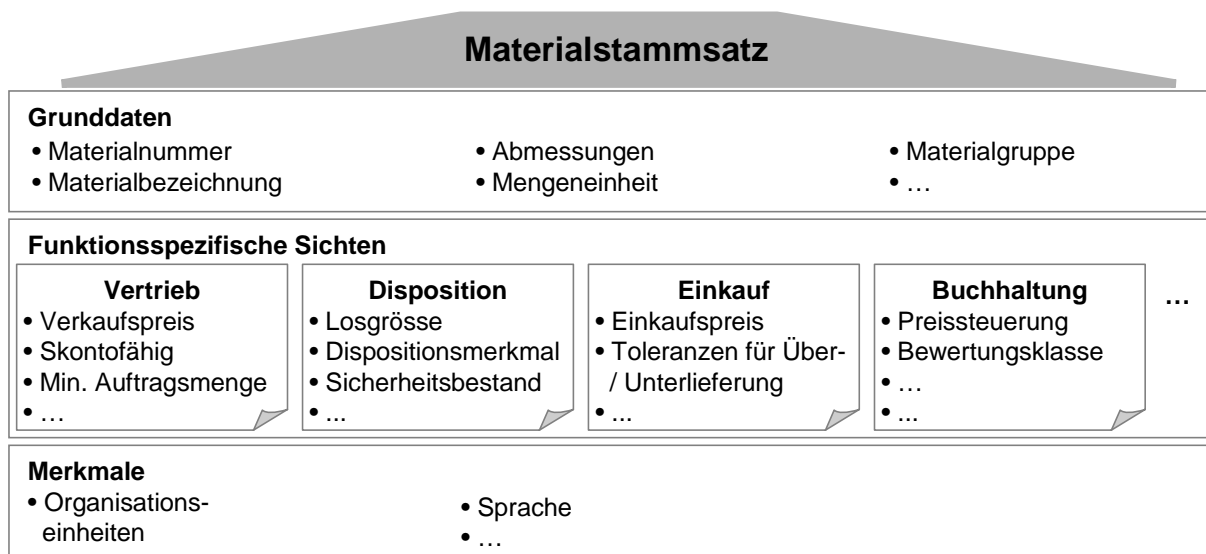


Abbildung 1: Beispiel für den Aufbau des Materialstammsatzes am Beispiel SAP R/3

<RN>Stamm-, Bewegungs- und Bestandsdaten</RN>

Stammdaten sind von Bewegungs- und Bestandsdaten abzugrenzen: Im Gegensatz zu den zustandsorientierten Stammdaten sind Bewegungsdaten abwicklungsorientiert und entstehen in Geschäftsvorfällen (z.B. Auftragseingang oder Materialentnahme). Bestandsdaten hingegen beschreiben die betriebliche Mengen- und Wertestruktur. So ist z.B. der verfügbare Lagerbestand eines Materials ein Bestandsdatum, das mit einer Warenbewegung geändert wird, während Materialnummer, -beschreibung und Mengeneinheiten als Stammdaten unverändert bleiben.

<RN>Informationslebenszyklus</RN>

Neben den verschiedenen Sichten spielt für die Bewirtschaftung von Stammdaten der Informationslebenszyklus eine wichtige Rolle. So entstehen erste Grunddaten zu einem Produkt bereits in der Produktentwicklung und werden später für die Produktion und schließlich den Vertrieb erweitert. Analog lassen sich für einen Kundenstammsatz verschiedene Lebenszyklusphasen festhalten: Der Interessent, der sich über Produkte informiert und ggf. eine Offerte verlangt, wird durch den Kauf zum Kunden und dann zum Nutzer eines Produkts, der durch eine Kundenservice-Einheit betreut wird.

Frage 1: Erläutern Sie am Beispiel „Kunde“ die Eigenschaften von Stammdaten.

2. Bedeutung des Stammdatenmanagements

<RN>Verwaltung und Aufgaben</RN>

Stammdatenmanagement beschränkt sich nicht allein auf die Informationsmodellierung und die Festlegung, in welchen Systemen diese gehalten werden, sondern umfasst vielmehr zahlreiche betriebswirtschaftliche und organisatorische Fragestellungen. Es besitzt also sowohl eine betriebswirtschaftlich-organisatorische als auch eine informationstechnische Kom-

ponente. Vor diesem Hintergrund beschäftigt sich das Stammdatenmanagement – als Bestandteil des gesamten betrieblichen Datenmanagements (vgl. Krcmar 2005) – vor allem mit folgenden Aufgaben:

- Beschreibung der Stammdatenstrategie, also der wesentlichen Ziele und der unternehmensweiten Richtlinien und Standards für den Umgang mit Stammdaten, mit deren Hilfe die Verfügbarkeit, Integrität und Sicherheit von Stammdaten gewährleistet wird,
- Festlegung der organisatorischen Verantwortung für Pflege und Erfassung der Stammdaten sowie für ihre Bereitstellung,
- Entwicklung eines Informationsmodells für sämtliche Stammdatenobjekte auf Unternehmensebene,
- Entwicklung und Pflege der Stammdatenarchitektur inkl. der Systeme zur Stammdatenhaltung und deren Verteilung.

<RN>Treibende Faktoren</RN>

Die Notwendigkeit für das Management von Stammdaten besteht, seitdem betriebliche Informationssysteme in Unternehmen eingesetzt werden. Für die Realisierung von Integrationseffekten ist die Verwendung eines einheitlichen Stammdatenmodells eine grundlegende Voraussetzung. Eine besondere Bedeutung erlangte das Stammdatenmanagement in den Neunziger Jahren im Zuge der unternehmensweiten Integration von Geschäftsprozessen und der Einführung von betrieblichen Standardanwendungssystemen (Enterprise Resource Planning Systems, ERP). Geschäftsprozessintegration ist nach wie vor ein treibender Faktor, doch darüber hinaus muss das Stammdatenmanagement heute zusätzlichen Erfordernissen Rechnung tragen. Die wichtigsten sind:

- Zwischenbetriebliche Vernetzung: Die Zusammenarbeit entlang der Wertschöpfungskette setzt gemeinsame und abgestimmte Datenbestände voraus. Einzelhändler und Konsumgüterhersteller können durch Synchronisation der Artikelstammdaten ihre Supply-Chain-Kosten um 1 bis 3 Prozent senken und gleichzeitig die „Speed-to-Shelf“, d.h. die benötigte Zeitspanne für die Einführung neuer Produkte, erheblich verbessern (vgl. Capgemini 2004).
- Gesetzliche Anforderungen: Regulatorische Auflagen schreiben zunehmend Transparenz und Rückverfolgbarkeit vor. Vorgaben wie der amerikanische Sarbanes-Oxley-Act sowie branchenspezifische Anforderungen, wie die EU-Verordnung zur Registrierung, Bewertung, Zulassung und Beschränkung chemischer Stoffe (REACH), sind nur mit unternehmensweit abgestimmten Stammdaten einzuhalten.
- Unternehmensfusionen und -übernahmen: Zusammenschlüsse und Übernahmen von Unternehmen erfordern, dass möglichst schnell entweder eine einheitliche oder zumindest eine vergleichbare Datenbasis zur Durchführung des operativen Geschäfts entsteht. Die gemeinsame Datenbasis muss sich mindestens auf die bilanzierungsrelevanten Kon-

solidierungsdaten erstrecken, kann – bei einer Verschmelzung der betrieblichen Aktivitäten – aber auch die vorhandenen Datenbestände vollständig umfassen.

- Kundenbeziehungsmanagement: Die Steigerung der Kundenzufriedenheit, die Nutzung von „Cross Selling“-Potenzialen sowie das frühzeitige Erkennen abgangsgefährdeter Kunden basieren auch auf einem effizienten und effektiven Management von Kundendaten.
- Bereitstellung von Managementinformationen: Fundierte Managemententscheidungen erfordern zusammenhängende und konsistente Daten auf allen Unternehmensebenen und über alle Geschäftsfelder hinweg. Der Aufbau eines Reportingsystems, z.B. einer Balanced Scorecard für das Gesamtunternehmen, bedingt daher konsistente Datenbestände.
- Informationstechnische Weiterentwicklungen: Die vorhandenen Stammdaten müssen beispielsweise bei einem Wechsel auf ein neues System-Release oder bei Migrationen von Altsystemen mindestens übernommen, besser sogar bereinigt werden. Auch ist bei einer Neuausrichtung der Anwendungslandschaft – z.B. im Zuge einer Systemkonsolidierung – die Qualität des Stammdatenmanagements zu gewährleisten. Häufig werden in derartigen Fällen zentrale Stammdatensysteme eingerichtet, aus denen die übrigen Anwendungssysteme mit Stammdaten beliefert werden.

<RN>Nutzenpotenziale</RN>

Ein gut funktionierendes Stammdatenmanagement generiert für sein Unternehmen betriebswirtschaftlichen Nutzen. Direkte Nutzenpotenziale entstehen durch eine Reduktion von administrativem Aufwand zur Erfassung, Pflege und Verteilung von Stammdaten sowie einer schnelleren Verfügbarkeit der Daten in den jeweiligen Anwendungssystemen. Als bedeutender sind jedoch die indirekten Nutzenpotenziale zu bewerten, die durch korrekte und aktuelle Stammdaten entstehen. Dazu gehören beispielsweise akkuratere Managementinformationen sowie Kosteneinsparungen und die Erhöhung der Prozessqualität (z.B. durch fehlerfreie und durchgängige Prozesse) (vgl. Russom 2006).

3. Stammdatenqualität

<RN>Definition von Stammdatenqualität</RN>

Grundsätzlich sind Daten dann von einer hohen Qualität, wenn sie den vom Nutzer festgelegten Zweck erfüllen. Insofern wird Datenqualität auch häufig mit dem Begriff „Fitness for Use“ beschrieben (vgl. PricewaterhouseCoopers 2005). Die konkretere Bestimmung der Qualität sollte verschiedene Kriterien berücksichtigen (vgl. Pipino et al. 2002):

- Vollständigkeit: Über das Vollständigkeitskriterium wird der Anteil an Nicht-Null-Werten eines bestimmten Datensatzes angegeben. In der Serienfertigung müssen beispielsweise vor Anlauf einer Serie sämtliche Attribute des Produkts (z.B. Abmessungen, Oberflächenbeschaffenheit) vorliegen, bevor mit der Fertigung begonnen werden kann.

- Aktualität: Die Aktualität beschreibt, wie zeitnah Datenwerte bei einer Änderung des realen Zustands angepasst werden. Anforderungen an Aktualität können durchaus unterschiedlich sein. So ist beispielsweise eine monatliche Aktualisierung der Mitarbeiterzahl eines Unternehmens für Berichtszwecke ausreichend, wohingegen die Aktualität von Produktdaten bei Just-in-Time-Produktion sekundengenau sein muss.
- Konsistenz: Konsistente Datenbestände zeichnen sich dadurch aus, dass sie frei von Widersprüchen sind. Beispielsweise sind Lieferantenstammdaten konsistent, wenn für jeden Lieferanten eine eindeutige Kennzeichnung existiert, wenn also keine Dubletten vorhanden sind.
- Zugänglichkeit: Der Umfang, in dem Daten für den Nutzer verfügbar bzw. auffindbar sind, wird durch das Kriterium der Zugänglichkeit bestimmt. Bei Online-Verkaufskanälen müssen Artikelinformationen wie Bilder oder Anwendungsbeispiele für den Konsumenten in Echtzeit zugänglich sein.
- Interpretierbarkeit: Dieses Kriterium beschreibt, inwieweit Daten in geeigneten Formaten vorliegen und dabei klare Regeln und Definition eingehalten werden. So müssen Daten beim Electronic Data Interchange (EDI) gewissen syntaktischen Vorgaben entsprechen, damit eine automatische Verarbeitung erfolgen kann. Eine Verletzung derartiger syntaktischer Regeln führt zu Fehlermeldungen in der EDI-Verarbeitung und erfordert zumeist eine manuelle Nachbearbeitung.

Die Verwendung der hier aufgeführten Kriterien objektiviert und konkretisiert den Begriff der Datenqualität und macht Datenqualität messbar. Erst über die Messung ist die Ableitung von Maßnahmen zur Verbesserung der Datenqualität möglich. Sie schafft die Voraussetzungen, um die in Abschnitt 2 dargestellten Anforderungen zu erfüllen.

Frage 2: Bei einem Unternehmen stehen Einkaufsverhandlungen an. Welche Qualitätskriterien sind in dieser Situation besonders wichtig in Bezug auf Lieferanten- und Materialstamm?

<RN>Ursachen mangelnder Stammdatenqualität</RN>

Aufgrund der steigenden betriebswirtschaftlichen Anforderungen wachsen sowohl die Bedeutung der Stammdatenqualität in Unternehmen als auch die Anstrengungen, Qualitätsprobleme zu beheben (vgl. Russom 2006). Ursachen für derartige Probleme sind häufig folgende Faktoren (vgl. PricewaterhouseCoopers 2005):

- Fehler bei der Dateneingabe: In manchen Geschäftsprozessen werden Daten manuell erfasst, z.B. bei der Aufnahme von Adressdaten durch Call-Center-Mitarbeiter oder bei der Online-Erfassung durch den Kunden selbst. Typische Fehler sind Rechtschreibfehler, die Eingabe von unpassenden Angaben in Muss-Felder, die nicht automatisch auf die Verwendung erlaubter Werte überprüft werden, oder die fehlerhafte Auswahl von Werten

aus vorgegebenen Listen (z.B. bei Mengeneinheiten die Verwechslung von Kilogramm und Gramm).

- Widersprüchliche Definitionen für übergreifend verwendete Datenobjekte: Häufig ist das Stammdatenmanagement in verschiedenen Geschäftsbereichen eines Unternehmens in der Historie unabhängig voneinander entstanden. Infolgedessen werden gleiche Datenobjekte bereichsweise verschieden interpretiert. Ein Beispiel aus der Automobilindustrie ist die mehrfache Definition des Datenobjekts „Getriebe“ als einerseits „Aggregat“ und andererseits „Komponente“.
- Fehler bei der Migration von Daten in neue Systeme: Bei der Einführung oder Zusammenführung von betrieblichen Informationssystemen (z.B. ERP-Systeme) müssen Daten von Altsystemen in die neue Systemumgebung überführt werden. Liegen die Daten nicht konsistent und syntaktisch einheitlich interpretierbar vor, kommt es bei der – zumeist automatisierten – Datenübernahme zu falschen Einträgen.
- Fehler bei der Integration von Daten aus mehreren Systemen oder aus externen Quellen aufgrund unterschiedlicher Datenmodelle: Häufig müssen für Berichtszwecke Daten aus unterschiedlichen Quellsystemen zusammengeführt und aggregiert werden. Besitzt ein Unternehmen beispielsweise weltweit drei verschiedene ERP-Systeme, möchte aber eine Analyse über die global beschafften Rohmaterialien durchführen, müssen die entsprechenden Materialstammdaten in einem einheitlichen Modell vorliegen oder mindestens in ein einheitliches Modell überführt werden können.

<RN>Aktueller Stand in der Praxis</RN>

Trotz ihrer Bedeutung wird die Stammdatenqualität von vielen Unternehmen noch vernachlässigt. So überprüfen von 400 befragten deutschen Unternehmen nur die Hälfte die Konsistenz ihrer Stammdaten regelmäßig (vgl. Quantz/Wichmann 2003).

4. Organisation des Stammdatenmanagements

<RN>Data Governance</RN>

Im betrieblichen Alltag werden stammdatenbezogene Aufgaben und Verantwortlichkeiten über unterschiedliche Geschäftsbereiche und Geschäftsprozesse verteilt wahrgenommen. Data-Governance-Konzepte setzen hier an und bilden einen Ordnungsrahmen, um trotz der verteilten Verantwortung die geforderte Verankerung im Unternehmen zu erreichen. Aufgrund des bereichs- und prozessübergreifenden Charakters wird Data Governance auch als „kollaborativer Prozess“ bezeichnet (vgl. PricewaterhouseCoopers 2005). Damit sind drei wesentliche Aufgaben verbunden:

- Identifikation der wesentlichen Strukturelemente und Steuerungsinstrumente im Stammdatenmanagement (Was?)
- Identifikation der betroffenen Rollen und Gremien im Unternehmen (Wer?)

- Zuordnung von Verantwortlichkeiten (Wie?)

Die wichtigsten Rollen und Gremien im Stammdatenmanagement sind (vgl. Dyché/Levy 2006):

- Das Data-Governance-Komitee ist auf Managementebene verantwortlich für die Entwicklung von Standards und Richtlinien, von Leitlinien für Stammdatenprojekte und für die Kommunikation der Ziele des Stammdatenmanagements im Unternehmen.
- Der Datenverantwortliche ist innerhalb eines bestimmten Geschäfts- oder Prozessbereichs für die Daten und ihre Qualität zuständig. Häufig ist der Datenverantwortliche auch mit der Umsetzung des Stammdatenmanagements gemäß der im Data-Governance-Komitee verabschiedeten Standards und Richtlinien betraut. In diesen Fällen wird für seine Rolle auch der Begriff „Data Steward“ verwendet.
- Der Prozessverantwortliche trägt die fachliche Verantwortung für einen Geschäftsprozess und die verwendeten Datenobjekte.
- Der Systemverantwortliche ist für das Informationssystem zuständig, in dem die Daten abgelegt sind.

<RN>Stammdatenprozesse</RN>

Neben dem übergreifenden Ordnungsrahmen, der durch Data Governance zu setzen ist, gehört zur Organisation des Stammdatenmanagements die Definition der operativen Stammdatenprozesse. Generisch sind dies die Anlage von Stammdaten, deren Pflege und ggf. Aktualisierung bei Änderungen und das Ausphasen bzw. Archivieren. Für das Management von Materialstammdaten lassen sich diese Prozesse wie folgt ausdrücken:

- Anlage: Die Auslöser für die Anlage eines neuen Materialstammdatensatzes können vielfältig sein und hängen häufig von der Art des Materials ab. Die Anlage eines neuen Fertigprodukts beispielsweise wird von der Entwicklungsabteilung angestoßen, während die Anlage eines Rohmaterials vom Einkauf ausgelöst wird. Die Anlage sämtlicher Attribute erfolgt sukzessive gemäß der Sichten durch die entsprechenden Prozessverantwortlichen.
- Pflege: Über den Informationslebenszyklus eines Materials ergeben sich Änderungen, die in den Stammsatz übernommen werden müssen. Auslöser dafür sind die Änderung der Herstellung eines Produkts, die Änderung der Verpackung oder die Erweiterung des Verkaufsgebiets.
- Ausphasen: Wenn Materialdaten nicht mehr benötigt werden, z.B. weil ein Artikel schon längere Zeit vom Markt genommen wurde, müssen sie deaktiviert werden. In der betrieblichen Praxis dürfen derartige Stammsätze nicht sofort gelöscht werden, weil die Daten noch in historischen Bewegungsdaten (Bestellungen, Fertigungsaufträge usw.) referenziert werden. Stattdessen werden die Stammdatensätze zunächst deaktiviert und erst gelöscht, wenn alle Transaktionsdaten abgearbeitet wurden.

Frage 3: Benennen Sie Prozessverantwortliche, die in die Anlage und die Pflege der Materialstammdaten eines Produktionsunternehmens involviert sind.

5. Stammdatenarchitekturen

<RN>Architekturansätze</RN>

Zur Umsetzung des Stammdatenmanagements in einer verteilten Systemlandschaft stehen vier idealtypische Architekturansätze zur Verfügung: Führendes Stammdatensystem, zentrales Stammdatensystem, Föderation über ein Verzeichnis (Registry) und Standards. Diese Ansätze lassen sich anhand der Dimensionen „Stammdatenmodell“ und „Stammdatenpflege und -haltung“ unterscheiden (vgl. Abb. 2):

- In einem harmonisierten Stammdatenmodell sind die Stammdatenattribute (z.B. Name und Feldlänge) und deren Inhalte (z.B. eindeutige Identifikatoren oder Wertebereiche) global, d.h. je nach Reichweite unternehmensweit oder unternehmensübergreifend, festgelegt. Bei einem nicht oder nur partiell harmonisierten Stammdatenmodell sind die Stammdaten in den einzelnen lokalen Systemen unterschiedlich ausgeprägt bspw. abhängig von Produktgruppen, Regionen oder Geschäftsbereichen.
- Bei Stammdatenpflege und -haltung unterscheidet man, inwieweit Stammdaten zentral angelegt und gespeichert werden. Bei der dezentralen Pflege in den lokalen Systemen werden Daten eventuell doppelt oder inkonsistent gespeichert. Die zentrale Erfassung und Pflege der Stammdaten sichert hingegen die Konsistenz von Stammdaten, da mit dem zentralen System ein „Single Point of Truth“ existiert. Allerdings ist ein zentraler Ansatz mit geringerer Flexibilität verbunden, da die Daten zentral vorhanden sein bzw. nach der Anlage im zentralen System in die lokalen Systeme verteilt werden müssen.

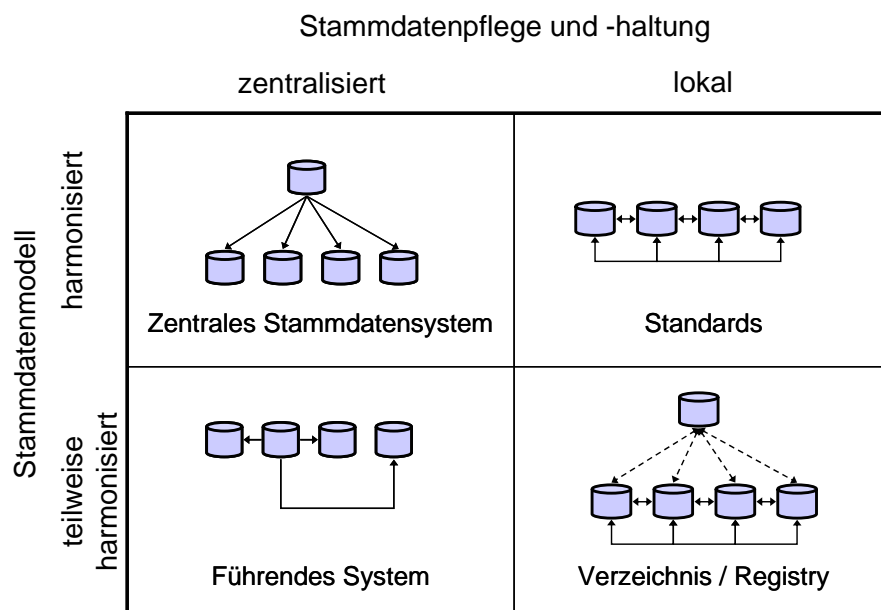


Abbildung 2: Einordnung der Architekturansätze

<RN>Asynchrone Steuerung, nicht ganz harmonisiert </RN>

„Führendes System“ ist das in der Praxis am häufigsten verwendete Verfahren für die Stammdatenverteilung: Eines der bestehenden Systeme wird als führendes System für eine Stammdatenklasse definiert und ist damit Ausgangspunkt für die Verteilung an die anderen Systeme, die asynchron entweder im Push- oder im Pull-Verfahren erfolgen kann. Das erstmalige Anlegen der Stammdaten geschieht im führenden System mit den dort vorhandenen Attributen. Da in dieser Variante kein vollständig harmonisiertes Datenmodell vorliegt, sind bei der Verteilung Datenergänzungen in den empfangenden Systemen sowie ein Mapping von Primärschlüsseln und Attributen auf das Datenformat des Zielsystems notwendig.

Beispiel: Eine Customer-Relationship-Anwendung wird als führendes System definiert. Bei der Eingabe der Stammdaten eines neuen Kunden (Name, Anschrift etc.) werden diese von der führenden Applikation an ein ERP-System und einen e-Shop verteilt. In den empfangenden Systemen sind Umschlüsselungen und Mappings notwendig.

<RN> Vollständig harmonisierte Steuerung </RN>

„Zentrales Stammdatensystem“ bedeutet Führung der Stammdaten in einem separaten Stammdatensystem, das diese an die lokalen Systeme verteilt. Auch bei diesem Ansatz findet die Erfassung und Pflege grundsätzlich in einem zentralen System statt, allerdings auf Basis eines harmonisierten Stammdatenmodells. Die Verteilung geschieht in der Regel asynchron (d.h. mit Verzögerung) und im Push-Verfahren. Obwohl mit neuen Ansätzen wie Serviceorientierten Architekturen (SOA) ein direkter Zugriff der Applikationen auf das Stammdatensystem theoretisch möglich wäre, werden die Stammdaten meist auch lokal gespeichert. Im Vergleich zu einem führenden System verfügt ein zentrales Stammdatensystem oft über zusätzliche Funktionalitäten und Workflow-Unterstützung für die Stammdatenprozesse.

Beispiel: Zentrale Stammdatensysteme werden derzeit durch neue Produkte von führenden Software-Anbietern wie SAP (Master Data Management) oder IBM (Product Information Center) forciert. Im überbetrieblichen Bereich sind Stammdatenpools wie SINFOS oder 1Sync Beispiele für zentrale Stammdatensysteme.

<RN>Nur logisch harmonisierte Koordination </RN>

„Standards“ führen nicht zu einer zentralen Speicherung und Verteilung von Stammdaten, es werden nur unternehmensweit einheitliche Strukturen definiert. Ein harmonisiertes Stammdatenmodell gewährleistet, dass Aufbau und Befüllung eines Stammdatensatzes über verschiedene Systeme hinweg gleich sind. Die Datenerfassung, -haltung und -pflege erfolgt jedoch dezentral in den lokalen Systemen. Die Standardisierung der globalen Attribute stellt sicher, dass einerseits ein Minimum an Attributen erfasst wird und diese andererseits in jedem System dieselbe Bedeutung haben. In der praktischen Umsetzung führt ein auf Stan-

dards basierender Ansatz häufig zu gewissen Inkonsistenzen im Datenbestand, z.B. Dubletten, da kein Stammdatenabgleich vorgesehen ist und kein „Single Point of Truth“ existiert.

Beispiel. Ein Konzern definiert ein so genanntes „Master Data Set“ für Produktdaten. Dieses Set enthält die als verbindlich definierten Attribute mit Beschreibungen (z.B. Codes), die für alle Konzerneinheiten Gültigkeit haben und in den lokalen Systemen umgesetzt werden.

<RN>Vollständige Koordination durch verteilte Einmalspeicherung</RN>

„Föderation über ein Verzeichnis“ (Registry) ist der Weg, ein übergreifendes Verzeichnis zu implementieren, das Zuordnungen der verschiedenen Stammdatensätze zu den verschiedenen Quellsystemen enthält. Benötigt beispielsweise ein System Daten zu einem bestimmten Kunden, so startet es eine Anfrage an die Registry und erhält eine Antwort, in welchem System die Daten zu dem jeweiligen Kunden abgelegt sind. In einem weiteren Schritt werden die Daten dann direkt aus dem entsprechenden System abgerufen. Die Datensätze werden dezentral in den lokalen Systemen angelegt, gepflegt und gehalten. Eine Datenverteilung gibt es in diesem Szenario nicht, das Verzeichnis initiiert lediglich den Datenzugriff. Änderungen in den Stammdatensätzen sind bei jedem Zugriff sofort verfügbar. Dieser Ansatz wird typischerweise bei sehr großen und verteilten Datenbeständen verwendet.

Beispiel. Die Global Registry der GS1 ist eine internationale Datenbank, in der hinterlegt wird, welcher Artikel in welchem Stammdatenpool gespeichert ist. Der Zugriff auf den konkreten Stammdatensatz findet dann im jeweiligen Stammdatenpool statt.

<RN>Vergleich der Architekturansätze</RN>

Tabelle 1 fasst die wesentlichen Eigenschaften der vier Ansätze zusammen.

Tabelle 1: Vergleich der Architekturansätze

	FÜHRENDES SYSTEM	ZENTRALES STAMMDATEN-SYSTEM	STANDARDS	FÖDERATION / VERZEICHNIS
DATENPFLEGE	Zentral	Zentral	Dezentral	Dezentral + Anlage in Registry
DATENHALTUNG	Mehrfach (führendes System und lokale Systeme)	Einmal (bei online Zugriff) oder mehrfach (bei Verteilung)	Einmal	Einmal
VERTEILUNG	Asynchron / i.d.R. Pull	Synchron oder asynchron / i.d.R. Push	Keine Verteilung notwendig	Keine Verteilung
HARMONISIERTER DATENBESTAND	Eingeschränkt gewährleistet	Gewährleistet durch globales Datenmodell	Gewährleistet durch globales Datenmodell	Nur teilweise gewährleistet
KONSISTENZ DES DATENBESTANDS ÜBER SYSTEMGRENZEN HINWEG	Gewährleistet, führendes System ist „Single Point of Truth“	Gewährleistet, Stammdaten-system ist „Single Point of Truth“	Nur teilweise gewährleistet, Redundanzen und Inkonsistenzen möglich	Nicht gewährleistet, Redundanzen und Inkonsistenzen möglich

AKTUALITÄT DES DATENBESTANDS IN DEN SYSTEMEN	Im führenden System hoch, jedoch Verzögerungen bei Verteilung in lokale Systeme möglich	Im Stammdaten-system hoch, jedoch Verzögerungen bei Verteilung in lokale Systeme möglich	hoch	hoch
---	---	--	------	------

Frage 4: Die Firma KinderGlück stellt Spielwaren her und vertreibt diese über verschiedene Vertriebskanäle an Spielwarenläden (u.a. Großhandel und direkter Online-Kanal). Im Rahmen einer CRM-Einführung sollen auch die Kundenstammdaten bereinigt werden. Welche Architekturvarianten eignen sich, wenn die Firma einen zentralen Ansatz verfolgen möchte? Welche Vor- und Nachteile ergeben sich?

Literaturempfehlungen

- Capgemini: Gute Gründe für globale Standards - Entwicklung eines Business Case für die globale Datensynchronisierung in Ihrem Unternehmen, Berlin 2004.
- Dyché, J./Levy, E.: Customer Data Integration, John Wiley & Sons, Inc., Hoboken 2006.
- Krcmar, H.: Informationsmanagement, Berlin 2005.
- Pipino, L. L./Lee, Y. W./Wang, R. W.: Data Quality Assessment, Communications of the ACM, Vol. 45 (2002), No. April, pp. 211-218.
- PricewaterhouseCoopers: Datenqualität als Grundlage für gute Managemententscheidungen. 2005
- Quantz, J./Wichmann, T.: E-Business-Standards in Deutschland, Berlecon Research, Berlin 2003.
- Russom, P.: Master Data Management: Consensus-Driven Data Definitions for Cross-Application Consistency, The Data Warehousing Institute, Renton, WA 2006.

Kontrollfragen und Antworten

Frage 1: Erläutern Sie am Beispiel „Kunde“ die Eigenschaften von Stammdaten.

Kundendaten gehören zu den wesentlichen Kerndaten eines Unternehmens, die sich nur selten ändern, z.B. im Falle einer Namens- oder Adressänderung. Ein Kundendatensatz besteht aus Grunddaten (z.B. Kundennummer, Name und Adresse) sowie verschiedenen Sichten: Für die Vertriebsicht sind z.B. die Zuordnung zu einer Kundengruppe und einer Vertriebsseinheit zentral, um daraus die Verkaufsprozesse und die Verkaufspreisermittlung zu steuern. Aus Sicht der Buchhaltung sind die Kreditorennummer sowie Angaben zur gewünschten Zahlungsweise relevant, während die Logistiksicht die Lieferadresse benötigt. Kundenstammdaten werden in verschiedensten Geschäftsvorfällen genutzt, z.B. bei der Erstellung eines Angebots, der Annahme eines Auftrags oder der Lieferung. Von zentraler Bedeutung ist, dass sämtliche Bereiche eines Unternehmens auf einen zentralen Stammsatz zugreifen, um Inkonsistenzen und redundante Datenhaltung zu vermeiden.

Frage 2: Bei einem Unternehmen stehen Einkaufsverhandlungen an. Welche Qualitätskriterien sind in dieser Situation besonders wichtig in Bezug auf Lieferanten- und Materialstamm? Bei Einkaufsverhandlungen ist es für das einkaufende Unternehmen sehr wichtig, dass zu jedem Lieferanten im Detail bekannt ist, wie hoch das Einkaufsvolumen in einer bestimmten Zeitperiode war. Je höher das Einkaufsvolumen ist, desto eher ist der Lieferant bereit, Preisnachlässe zu gewähren. Deshalb müssen die Daten des Lieferanten- und des Materialstamms konsistent sein, also keine Dubletten existieren, die eine Berechnung des gesamten Volumens verhindern. Außerdem müssen die Daten aktuell sein, denn veraltete Materialien, die nur in vergangenen Zeitperioden von einem Lieferanten bezogen wurden, dürfen nicht in aktuelle Auswertungen einbezogen werden.

Frage 3: Benennen Sie Prozessverantwortliche, die in die Anlage und die Pflege der Materialstammdaten eines Produktionsunternehmens involviert sind.

Entlang des Informationslebenszyklus sind unterschiedliche Prozessverantwortliche in das Management des Materialstamms involviert. Prozessverantwortliche aus der Produktentwicklung sind z.B. für die Anlage von Grunddaten verantwortlich und erstellen eine Stückliste aus Sicht der Konstruktion. Der Prozessverantwortliche aus dem Marketing ergänzt Produktbeschreibungen und Bilder für die Kunden, während der Prozessverantwortliche aus der Produktion die Produktionsstückliste anlegt und Daten zur Arbeitsvorbereitung unterhält. Der Prozessverantwortliche aus Vertrieb und Disposition erfasst Angaben zur Verpackung und zum Versand des Produkts

Frage 4: Die Firma KinderGlück stellt Spielwaren her und vertreibt diese über verschiedene Vertriebskanäle an Spielwarenläden (u.a. Großhandel und direkter Online-Kanal). Im Rahmen einer CRM-Einführung sollen auch die Kundenstammdaten bereinigt werden. Welche Architekturvarianten eignen sich, wenn die Firma einen zentralen Ansatz verfolgen möchte? Welche Vor- und Nachteile ergeben sich?

Da ein zentraler Ansatz gewünscht ist, kommen nur zwei Alternativen in Frage: Das Unternehmen kann entweder eines der bestehenden Informationssysteme als führendes System definieren oder ein separates Stammdatensystem aufbauen. Im ersten Fall könnte z.B. das neue CRM-System zum führenden System werden, während im zweiten Fall ein weiteres System aufgebaut werden müsste. Zentrale Ansätze haben einige Vorteile, da Daten jeweils zentral gepflegt werden und damit Konsistenz und Redundanzfreiheit des Datenbestands einfacher gewährleistet werden. Allerdings können sich Nachteile ergeben, weil die Daten in der Regel in die lokalen Systeme verteilt und mehrfach gespeichert werden. Dadurch sind Aktualisierungen ggf. nicht sofort verfügbar.