



# Enterprise Information Integration

## Die tragende Säule einer modernen Informationsmanagement-Architektur

Unternehmen, die zukünftige Kundenbedürfnisse antizipieren und sich dadurch neuen Produkten und Märkten öffnen und sie sogar mitgestalten, müssen verstärkt auf das Management ihrer eigenen Kompetenzen achten. Als Grundlage eines nachhaltigen Unternehmenserfolgs müssen sie den Aufbau von Kernkompetenzen betreiben. Informationen und Know-how werden zum erfolgsentscheidenden Produktionsfaktor und Differenzierungsmerkmal am Markt.

Erfolg kann dieser Ansatz nur haben, wenn die geschäftskritische Ressource „Information“ professionell verwaltet und entlang der Wertschöpfungskette optimal integriert wird.

Daten sind nur dann wirklich wertvoll, wenn sie effizient genutzt werden. Doch allzu oft machen es die Art der Speicherung sowie die komplexen Prozesse beim Verteilen und Zugreifen auf Daten für den Endanwender nahezu unmöglich, an die gewünschten Informationen zu gelangen.

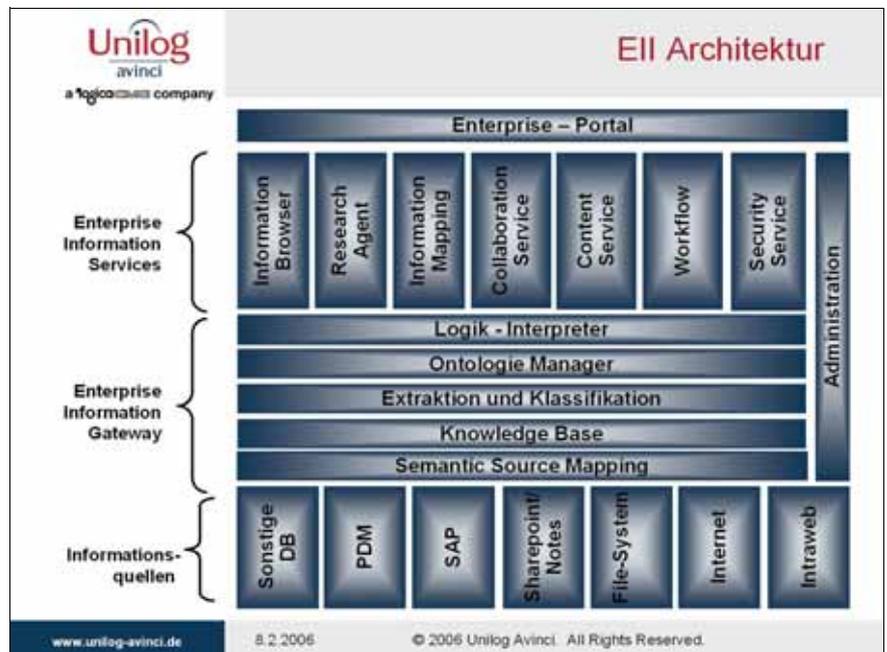
Sind Informationen nicht zugänglich, unpräzise oder ungeordnet, werden sie schnell nutzlos oder – was noch schlimmer ist – führen zu falschen Entscheidungen, unzufriedenen Kunden und letztendlich zu einem Verlust von Wettbewerbsfähigkeit.

Dieses hat nicht nur Auswirkungen auf diejenigen Prozesse und Werkzeuge, die im Rahmen eines Informationsmanagements die Ressource Information im Unternehmen identifizieren, generieren, organisieren, verwalten und kontinuierlich weiterentwickeln, sondern auch auf die Gestaltung und kontinuierliche Steuerung der operativen Geschäftsprozesse, die diese Informationen für die Erstellung von Produkten und Dienstleistungen nutzen und entwickeln.

### Zielarchitektur für ein ganzheitliches Informationsmanagement

Das Rückgrat einer durchgängigen Informationsmanagement-Umgebung wird durch eine Enterprise-Content-Management-Service-Architektur gestellt, welche eine Synthese aus zwei wesentlichen IT-Trends – BPM und SOA – darstellt.

Business Process Management (BPM) extrahiert die Prozesslogik von den Anwendungen und ermöglicht Prozesse sichtbar und kontrollierbar zu machen. BPM schließt die Lücke zwischen Fachbereich und IT.



Service Oriented Architecture (SOA) ist eine Architektur, die auf der Bereitstellung und Nutzung von Services für Anwendungen basiert. Services unterstützen eine fachliche Funktion, sind wiederverwendbar in verschiedenem Kontext/Kanal und Prozessen und sind ein Konstrukt und keine Technologie (im Unterschied zum Beispiel zu Web Services). Services werden über Prozesse „orchestriert“ und in verschiedenen Prozessen wieder verwendet.

Die einzelnen Basistechnologien für Archivierung, Dokumenten-, Workflow-, Content- und Knowledge Management liefern die systemtechnisch bereitgestellten Dienste im Sinne derartiger Services. Diese werden entsprechend den Erfordernissen innerhalb der Kernprozesse individuell zur Verfügung gestellt. Content Repositories und deren Erfassungs- und Retrievalservices ermöglichen die prozessspezifische Aufbereitung und Bereitstellung der Informationen.

So ungefähr lässt sich eine heute häufig anzutreffende Vision eines Unternehmens für ein ganzheitliches Informationsmanagement beschreiben. Allein die unternehmensintern gewachsenen Strukturen, Plattformen und Prozesse sowie die am Markt verfügbaren Systeme und Integrationsstandards setzen die restriktiven Rahmenbedingungen für dessen Realisierung. Kaum ein Hersteller ist derzeit in der Lage, ein derartiges technologisches Full-Service-Angebot zu liefern. Der Implementierungs-, Integrations- und Migrationsaufwand ist vergleichsweise hoch und manchmal prohibitiv.

Ein typisches Unternehmen verfügt über eine Vielzahl an Datenbanken und darauf aufbauenden Anwendungen. Die heterogenen Systeme sind oftmals das Resultat von Fusionen, einer Fülle von abteilungsspezifischen Anwendungen und die Einführung von Insellösungen, die zur Erfüllung eines spezifischen Zwecks implementiert wurden. Es gibt Banken, die mehr als 40 verschiedene Call-Center-Anwendungen im Einsatz haben, Unternehmen mit über 1.000 Datenbanken und Hersteller im produzierenden Gewerbe mit mehr als 2.000 CAD/CAM-Systemen. Diese Systeme beinhalten wertvolle Informationen und sind geeignet dafür, ganz spezielle Aufgaben zu übernehmen. Unglücklicherweise können diese Informationen nicht ohne erheblichen Aufwand für andere Zwecke nutzbar gemacht werden. Wenn neue Anwendungen hinzukommen, die Anforderungen an Funktionalitäten sich ändern oder Datenstrukturen angepasst werden müssen, ist dies üblicherweise mit einem erheblichen Aufwand verbunden.

### Information Integration – das Problem

Die übliche Vorgehensweise von EAI-Lösungen zur Integration von Datenquellen ist die Verwendung von so genannten „Field-to-Field-Mappings“. Dabei werden die Schemata zweier Datenquellen importiert und die einzelnen Felder aufeinander abgebildet. Mit einfachen Regeln können Felder aufgeteilt oder verknüpft werden oder einfache Umwandlungen erreicht werden.

---

Die Umsetzung hat sich jedoch als recht mühsam erwiesen:

- o da „Field-to-Field-Mappings“ nicht skalieren. Die Anzahl an Mappings wächst exponential mit jeder neuen Datenquelle und Anwendung. Wartung und Weiterentwicklung werden zum Problem, da jede Änderung eines Schemas einer Datenquelle eine Anpassung sämtlicher darauf zugreifender Anwendungen und Mappings erforderlich macht.
- o Unternehmen stellen oft fest, dass die Erstellung von korrekten Mappings ein schwieriges Unterfangen ist. Es setzt voraus, dass die für das Mapping verantwortliche Person beide Datenquellen im Detail kennt oder zumindest versteht, was vor allem bei notwendigen Änderungen schier unmöglich ist. Folglich werden hierbei viele Fehler gemacht.
- o Das Mapping zwischen zwei Datenquellen, die unterschiedlichen Entwurfs-Paradigmen folgen, ist extrem komplex. Es existieren verschiedene Möglichkeiten, ein Schema zu entwerfen. Hierarchische Strukturen treffen auf relationale Datenbanken, flache Strukturen auf stark verschachtelte, normalisierte Schemas auf weniger normalisierte.
- o Die Mappings erlauben weder die Abbildung der Anwendungslogik noch die Abbildung der bestehenden Komplexität von Abhängigkeiten. Dieses notwendige Wissen, um Prozesse zu optimieren, bleibt fest in der Anwendung programmiert und kann nicht automatisiert anderen Anwendungen zur Verfügung gestellt werden.

### Die Lösung – Enterprise Information Integration durch Semantische Integration

Semantische Technologien ermöglichen einen völlig neuen Ansatz zur Integration von Anwendungen und Daten.

Vor der Realisierung von Mappings, wird hierbei ein Modell (Ontologie) erstellt, welches einen ausgewählten Geschäftsbereich definiert. Ein Ausgangspunkt für eine Ontologie kann beispielsweise der Import eines Datenbank-Schemas sein. Das objektorientierte Modell wird in einer Sprache zur Wissensrepräsentation (z. B. F-Logic, OXML, RDF, Daml+Oil, OWL) ausgedrückt und enthält Konzepte, Attribute, Relationen zwischen Konzepten und im Gegensatz zu anderen Repräsentierungsformen (z. B. Thesauri, TopicMaps) Regeln, die einem wohl definierten Logikformalismus folgen. Das so entstandene Wissensmodell wird mit den vorhandenen Datenquellen, wie beispielsweise Datenbanken, XML-Schemata oder auch Suchmaschinen verknüpft.

Dieser Ansatz löst eine Reihe von Wartungs- und Entwicklungsproblemen sowie Problemen in der Schema-Kompatibilität.

Durch die Verwaltung des Wissens in einer getrennten Schicht für den Systemeinsatz ermöglichen Ontologien die Kommunikation zwischen Systemen, unabhängig von individueller Technologie, Architektur und der Anwendungen. Als gemeinsames Modell ermöglichen sie die einheitliche Verwaltung von unternehmensweiten Datenquellen.

Der entscheidende Vorteil dieser Methodik ist deren Optimierung hin auf die Abbildung von Relationen zwischen Objekten sowie generische als auch spezifische Regeln. Damit kann die beschriebene Abbildungsproblematik drastisch vereinfacht werden. Semantische Ansätze vereinen somit eine Reihe von Vorteilen in der Integration:

- o Verständliche Darstellung in lesbaren Strukturen: Eine wesentliche Abgrenzung gegenüber z. B. einem konzeptuellen Datenmodell im Datenbankbereich ist, dass (i) sowohl Nomenklatur als auch Strukturierung in einem sozialen Prozess erstellt wurden und (ii) beides nach außen sichtbar wird. Somit ist eine hohe Akzeptanz darüber im Anwendungsgebiet sichergestellt. Das Modell und die darin definierten Relationen sind für den Anwender verständlich, nachvollziehbar und somit einfacher wartbar.
- o Wartungskosten können nachhaltig durch die Trennung von Struktur, Daten, und Anwendungen gesenkt werden. Änderungen an der Anwendung können somit vollzogen werden, ohne dass die Datenschicht umstrukturiert werden muss. Gleiches gilt für den umgekehrten Fall.
- o Flexible, dynamische und inhaltliche Integration. Mittels Semantischer Modelle lässt sich ein Datenmodell erstellen, in dem unterschiedliche Strukturen (z. B. relationale und hierarchische, aber auch stark und schwach strukturierte) kombiniert und gleichzeitig durch Regeln Namens-, Format- und Strukturkonflikte aufgelöst werden können. Über vernetzte Ontologien lassen sich darüber hinaus unterschiedliche Sichten auf Informationen abbilden.

### Zusammenfassung

Im Laufe der letzten Jahre hat sich der Schwerpunkt des IT-Managements verschoben: Ging es früher vorrangig um die Verwaltung der Technologie selbst, gilt es heute, mit IT einen Wertbeitrag für das Unternehmen zu leisten. Folglich bemühen sich viele Unternehmen um die Unterstützung ihres Geschäftsmodells durch IT-Lösungen und betrachten ihre interne IT-Organisation – oder jeden anderen Technologiepartner – nunmehr als „Service Provider“, als Dienstleister, der dem Unternehmen einen Nutzen bieten muss.

Gleichzeitig steigt die Komplexität der heutigen Unternehmens-IT mit jeder neuen Anwendung und mit jeder Weiterentwicklung bestehender Systeme in ungebremster Weise an. Und damit ist genau dieser Sachverhalt als Auslöser für steigende Betriebskosten, mangelnde Flexibilität und das Scheitern ambitionierter IT-Projekte bekannt.

Informationsintegration und -migration gehören in diesem Spannungsfeld zweifellos zu den Komplexitäts- und Kostentreibern. Informationstechnologie kann jedoch nur zu einem modernen Dienstleister avancieren wenn sie ihre Kosten minimiert und gleichzeitig die „neue Komplexität“ beherrscht.

Dieses wiederum ist nur erfolgreich, wenn Enterprise Information Integration als eine wesentliche tragende Säule einer serviceorientierten Unternehmensarchitektur verstanden wird und diese gleichzeitig weniger eine technische, als vielmehr eine Managementdisziplin darstellt.

Durch den Einsatz von semantischen Methoden und Technologien zur Beschreibung komplexer Zusammenhänge und deren Fähigkeit, heterogene Datenstrukturen zu integrieren, wird EII zur Semantic-Business-Integration-Schicht einer globalen Unternehmensarchitektur.

Enterprise Information Integration ist somit das semantische Netzwerk entlang einer durchgängigen Supply Chain.

### Ulf Freiberg,

Stellvertretender Vorstandsvorsitzender  
VOI e.V.

### Quellen:

„Dokumenten-Management – vom Archiv zum Enterprise Content Management“, Code of Practice: VOI e.V., Prof. Dr. Jürgen Angele  
„Semantic Business Integration – Speed up your Processes“, Andreas Nierlich, ontoprise GmbH,  
„White Paper Knowledge Relationship Management“, Unilog Avinci a LogicaCMG company