

Metadaten, Application Understanding und Software Migration

Clauß-Peter Hahn
ASG GmbH & Co. KG
Straßburger Straße 34
09120 Chemnitz
clauss-peter.hahn@asg.com

Was sind Metadaten?

Die immer wieder genutzte Beschreibung ist. „Daten die Daten beschreiben“ ist zwar wahr aber zu ungenau. In Wikipedia findet man eine etwas genauere Beschreibung: „Als Metadaten oder Metainformationen bezeichnet man allgemein Daten, die Informationen über andere Daten enthalten. Bei den beschriebenen Daten handelt es sich oft um größere Datensammlungen (Dokumente) wie Bücher, Datenbanken oder Dateien. So werden auch Angaben von Eigenschaften eines Objektes (beispielsweise Personennamen) als Metadaten bezeichnet.“ (<http://de.wikipedia.org/wiki/Metadaten>)

Diese Beschreibung zeigt sehr deutlich ein Problem der Verwendung des Begriffes Metadaten auf. Mit Metadaten beschreibt man Zusatzinformationen zu Daten, wie es in Bibliothekskatalogen (Dublin Core) oder bei Zusatzinformationen zu MP3 Daten (ID3) üblich ist, innerhalb der IT werden die physische und logische Struktur von Informationen wie in Datenbankkatalogen, im Datenfluss von IT Anwendungen oder in Schema von XML Daten und deren Beziehungen als Metadaten bezeichnet. Der Aspekt der Beziehungen wird leider außerhalb des reinen Metadaten Management ungenügend oder gar nicht betrachtet.

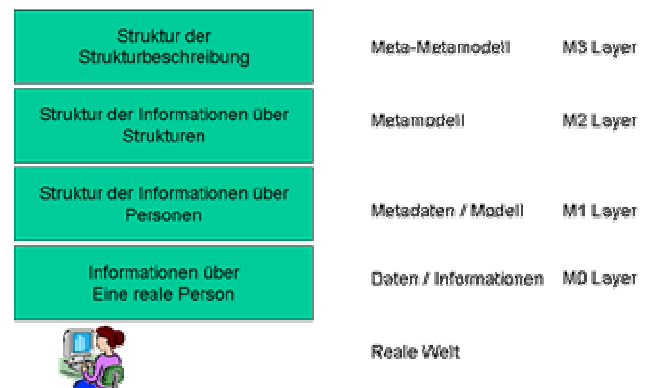
Aus der Sicht des Autors dienen Metadaten zur Beschreibung

- der Struktur von Informationen
- der Verwendung von Informationen
- der Klassifikation von Informationen
- der Beziehungen zu anderen Metadaten
- der Beziehungen zu Business Daten

Das sehr bekannte 4-Schichtenmodell¹, genutzt in MOF Version 1², erlaubt eine sehr gute Veranschaulichung der

¹ Eine genauere Analyse dieses Schichtenmodells führt unter Umständen dazu, dass eine 4-Schichten Architektur nicht hinreichend ist, was zu einer rekursiven Definition von Metadaten führt.

verschiedenen Metalayer und dient zur Definition von Begriffen im Umgang mit Metadaten:



Jeder Layer ist eine Sprache zur Beschreibung des nächst niedrigen Layers.

Auffinden von Metadaten

Metadaten können aus den unterschiedlichsten Dokumenten, welche im Software Entwicklungs- und Verwaltungsprozess entstehen gewonnen und zueinander in Beziehung gestellt werden.

Mögliche Metadaten Quellen sind:

- Datenbank Kataloge
- Sourcecode
- Generierungswerkzeuge
- Entwicklungswerkzeuge
- Business Informationen
- Data Warehouse Tools
- usw

² <http://www.omg.org/cgi-bin/apps/doc?formal/00-04-03.pdf>

Speicherung und Analyse von Metadaten

Metadaten werden in Repositories, einem Datenspeicher deren Struktur auf einem Metameta-Modell basiert, abgelegt. Repository sollen folgende Funktionalität anbieten:

- Speicherung von Metadaten basierend auf beliebig vielen Metamodellen
- Möglichkeit der dynamische Änderung von Metamodellen
- Strukturierung und Partitionierung von Metadaten
- Versionierung von Metadaten und Beziehungen
- Persistentes und dynamisches Bilden von Konfigurationen
- Paralleler Zugriff auf verschiedene Versionen bzw. Konfigurationen
- Unterstützung von Prozessen

Repository erlauben die einfache Auswertung komplexer Netzwerkstrukturen und deren Präsentation. Abhängig von der Art des Repository (Enterprise, Applikation, Tool) ist auch die Granularität der Metadaten und der Detailgrad der Auswertungen, nicht jedoch die Komplexität der Struktur der Metadaten selbst.

Auswertungen der Wirkung von Änderungen, und das Verständnis der gesamten IT Landschaft im Zusammenhang einschließlich der komplexen Zusammenhänge werden unterstützt und häufig erst möglich gemacht.

Da Repository sowohl die technischen als auch die Business Strukturen enthalten können auch die Beziehungen zwischen diesen Informationen ausgewertet und dargestellt werden.

Repository basierte Lösungen

Die in einem Repository gespeicherten Metadaten können in unterschiedlichen Abwendungen genutzt werden, dies sind:

- Data Warehouse
- Daten Management & Integration
- Enterprise Architektur
- Data Dictionary
- Enterprise Resource Planung
- Application Discovery & Understanding
- Application Portfolio Management
- Service Oriented Architecture

Software Migration

Software Migration wird allgemein als 4-stufiger Prozess dargestellt, wobei die Stufen 2 und 3 in der Reihenfolge getauscht werden können.

1. Understanding
 - Verstehen der Applikation und deren Einbettung in die firmenweite IT Landschaft
2. Remediation
 - Restrukturierung und Sanierung der existierenden Applikation
 - Entfernen toter Komponenten
 - Isolation der Business Logik/Rules als Komponenten
3. Migration
 - Transformieren der Applikation
 - Definition neuer Interfaces
 - Bilden von Objekten (der existierende Code ist nicht OO)
4. Maintenance
 - Pflege der Komponenten
 - Re-implementierung von Komponenten aus technischen oder Business Gründen

Die Komplexität des Migrationprozesses erhöht sich potenziell, da (1) das existierende System weiterhin gepflegt werden muss und diese Modifikationen auch migriert werden müssen und (2) existierende und migrierte (Teil-)Applikationen parallel weiter existieren müssen. Für einen sehr langen Zeitraum wird eine große Menge existierenden Codes entweder direkt oder als Service weiter in migrierten Applikationen vorhanden sein. Softwremigration ist die Kunst, eine Autobahn ohne Sperrungen, Geschwindigkeitsbegrenzungen und Behinderungen vollständig zu erneuern.

Dieser Prozess erfordert eine Verwaltung, Pflege und Analyse aller Schnittstellen und Beziehungen innerhalb eines zentralen Tools.

Repository stellen das geeignete Mittel zur begleitenden Dokumentation von Software Migration Projekten dar. Sie bieten alle notwendigen Analyse und Metadaten Auswertungswerkzeuge um existierende Applikation und deren Beziehungen untereinander zu verstehen. Die Möglichkeit die Beziehungen zwischen Algorithmen und den zugehörigen Business Prozessen explizit zu dokumentieren und in parallelen Konfigurationen zu halten, erlaubt die vollständige Dokumentation jeden Projektschrittes. Die Nutzung von Schnittstellen und die Auswirkungen von Änderungen kann analysiert werden. Die Data Dictionary Funktionalität erlaubt das Generieren von Datendefinitionen in unterschiedlichen Zielsprachen basierten auf einer zentralen Beschreibung.