

Christian GÜthlein

Metadatenmanagement

Erfolgsfaktor für ein Data Warehouse

Diplomarbeit

Bibliografische Information der Deutschen Nationalbibliothek:

Bibliografische Information der Deutschen Nationalbibliothek: Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de/> abrufbar.

Dieses Werk sowie alle darin enthaltenen einzelnen Beiträge und Abbildungen sind urheberrechtlich geschützt. Jede Verwertung, die nicht ausdrücklich vom Urheberrechtsschutz zugelassen ist, bedarf der vorherigen Zustimmung des Verlanges. Das gilt insbesondere für Vervielfältigungen, Bearbeitungen, Übersetzungen, Mikroverfilmungen, Auswertungen durch Datenbanken und für die Einspeicherung und Verarbeitung in elektronische Systeme. Alle Rechte, auch die des auszugsweisen Nachdrucks, der fotomechanischen Wiedergabe (einschließlich Mikrokopie) sowie der Auswertung durch Datenbanken oder ähnliche Einrichtungen, vorbehalten.

Copyright © 2000 Diplom.de
ISBN: 9783832445850

Christian G thlein

Metadatenmanagement

Erfolgsfaktor f r ein Data Warehouse

Christian GÜthlein

Metadatenmanagement

Erfolgsfaktor für ein Data Warehouse

Diplomarbeit

an der Bayerischen Julius-Maximilians-Universität Würzburg

Lehrstuhl für Betriebswirtschaftslehre und Wirtschaftsinformatik

November 2000 Abgabe



Diplomica GmbH _____
Hermannstal 119k _____
22119 Hamburg _____

Fon: 040 / 655 99 20 _____
Fax: 040 / 655 99 222 _____

agentur@diplom.de _____
www.diplom.de _____

ID 4585

Güthlein, Christian: Metadatenmanagement: Erfolgsfaktor für ein Data Warehouse / Christian

Güthlein - Hamburg: Diplomica GmbH, 2001

Zugl.: Würzburg, Universität, Diplom, 2000

Dieses Werk ist urheberrechtlich geschützt. Die dadurch begründeten Rechte, insbesondere die der Übersetzung, des Nachdrucks, des Vortrags, der Entnahme von Abbildungen und Tabellen, der Funksendung, der Mikroverfilmung oder der Vervielfältigung auf anderen Wegen und der Speicherung in Datenverarbeitungsanlagen, bleiben, auch bei nur auszugsweiser Verwertung, vorbehalten. Eine Vervielfältigung dieses Werkes oder von Teilen dieses Werkes ist auch im Einzelfall nur in den Grenzen der gesetzlichen Bestimmungen des Urheberrechtsgesetzes der Bundesrepublik Deutschland in der jeweils geltenden Fassung zulässig. Sie ist grundsätzlich vergütungspflichtig. Zuwiderhandlungen unterliegen den Strafbestimmungen des Urheberrechtes.

Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Werk berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften.

Die Informationen in diesem Werk wurden mit Sorgfalt erarbeitet. Dennoch können Fehler nicht vollständig ausgeschlossen werden, und die Diplomarbeiten Agentur, die Autoren oder Übersetzer übernehmen keine juristische Verantwortung oder irgendeine Haftung für evtl. verbliebene fehlerhafte Angaben und deren Folgen.

Diplomica GmbH

<http://www.diplom.de>, Hamburg 2001

Printed in Germany



Wissensquellen gewinnbringend nutzen

Qualität, Praxisrelevanz und Aktualität zeichnen unsere Studien aus. Wir bieten Ihnen im Auftrag unserer Autorinnen und Autoren Wirtschaftsstudien und wissenschaftliche Abschlussarbeiten – Dissertationen, Diplomarbeiten, Masterarbeiten, Staatsexamensarbeiten und Studienarbeiten zum Kauf. Sie wurden an deutschen Universitäten, Fachhochschulen, Akademien oder vergleichbaren Institutionen der Europäischen Union geschrieben. Der Notendurchschnitt liegt bei 1,5.

Wettbewerbsvorteile verschaffen – Vergleichen Sie den Preis unserer Studien mit den Honoraren externer Berater. Um dieses Wissen selbst zusammenzutragen, müssten Sie viel Zeit und Geld aufbringen.

<http://www.diplom.de> bietet Ihnen unser vollständiges Lieferprogramm mit mehreren tausend Studien im Internet. Neben dem Online-Katalog und der Online-Suchmaschine für Ihre Recherche steht Ihnen auch eine Online-Bestellfunktion zur Verfügung. Inhaltliche Zusammenfassungen und Inhaltsverzeichnisse zu jeder Studie sind im Internet einsehbar.

Individueller Service – Gerne senden wir Ihnen auch unseren Papierkatalog zu. Bitte fordern Sie Ihr individuelles Exemplar bei uns an. Für Fragen, Anregungen und individuelle Anfragen stehen wir Ihnen gerne zur Verfügung. Wir freuen uns auf eine gute Zusammenarbeit.

Ihr Team der Diplomarbeiten Agentur

Diplomica GmbH _____
Hermannstal 119k _____
22119 Hamburg _____

Fon: 040 / 655 99 20 _____
Fax: 040 / 655 99 222 _____

agentur@diplom.de _____
www.diplom.de _____

Inhaltsverzeichnis

1	Problematik.....	1
2	Metadaten	3
2.1	Begriffsdefinition.....	3
2.2	Aufgabenbereich.....	5
2.3	Anwendungsfelder	8
2.4	Metadatenkategorien	9
2.4.1	Administrative Metadaten	9
2.4.2	Berechtigungs-Metadaten.....	9
2.4.3	Metadaten zu strukturellen Aspekten.....	10
2.4.4	Metadaten zum Kontext und Inhalt.....	10
2.4.5	Metadaten zur Nutzungs- und Wirkungsgeschichte	11
3	Data Warehouse.....	11
3.1	Definition.....	12
3.2	Aufbau	13
3.3	Datenintegration.....	15
3.4	Abfrage.....	16
3.5	Erfolgsfaktoren für ein Data Warehouse.....	18
4	Management von Metadaten.....	19
4.1	Metamodell.....	19
4.2	Metadatenarchitektur.....	20
4.3	Aufgabe	22
4.4	Metadaten-Repository	24
4.4.1	Definition.....	24
4.4.2	Zentrales Repository	25
4.4.3	Dezentrales Repository.....	26
4.4.4	Metadatenfluss	27
4.4.5	Einsatzbereich	30
4.5	Metadatenstruktur	30
4.5.1	Back Room Metadaten	31
4.5.2	Front Room Metadaten.....	32
4.5.3	Aufgabenbereich	34

4.6	Metadatenaustausch.....	35
4.7	Metadatenintegration.....	38
5	Standardisierungsbestrebungen.....	39
5.1	Bewertungskriterien	40
5.2	Object Management Group (OMG).....	42
5.2.1	Object Management Architecture (OMA).....	43
5.2.2	UML	43
5.2.3	Meta Objekt Facility (MOF).....	44
5.2.3.1	MOF-Modell.....	45
5.2.3.2	MOF IDL Mapping.....	47
5.2.3.3	MOF Schnittstellen	48
5.2.4	XML Metadata Interchange Format (XMI).....	49
5.2.4.1	Aufgabe	49
5.2.4.2	Austausch.....	51
5.3	Common Warehouse Metamodel.....	54
5.3.1	Aufbau	55
5.3.2	CWM-Metamodelle.....	57
5.4	Meta Data Coalition (MDC)	58
5.4.1	Open Information Model (OIM)	59
5.4.2	XML Interchange Format (XIF).....	61
5.4.3	Aufbau eines OIM Modells.....	61
5.4.3.1	Analyse und Design Modell	62
5.4.3.2	Datenbank und Data Warehousing Modell	62
5.4.3.3	Objekt und Komponenten Modell.....	63
5.4.3.4	Knowledge Management Modell.....	63
5.4.3.5	Business Engineering Modell.....	63
5.4.4	Aufbau des Modells.....	63
5.4.4.1	Datenbank Modell.....	64
5.4.4.2	Data Warehouse Modell.....	64
5.5	Spezifische Metadatenaustausch-Formatstandards	65
5.6	Vergleich von OIM und CWM	66
5.7	Bewertung.....	69
6	Anwendungen.....	70
6.1	Oracle Warehouse Builder	70
6.2	Microsoft Repository.....	72

6.3	Platinum Repository.....	73
6.3.1	Platinum Repository Architektur	74
6.3.2	Schnittstellen.....	75
6.3.3	Einsatz der Standardspezifikationen	76
6.4	Informatica.....	77
6.4.1	Metadaten austausch	77
6.4.2	Schnittstelle.....	78
7	Fazit	79
	Quellenverzeichnis	81
	Abkürzungsverzeichnis.....	88

1 Problematik

In dieser Arbeit wird die Grundproblematik und das geforderte Anforderungspotential von Metadaten und deren Verwaltung für den effizienten Aufbau eines leistungsstarken Data Warehouses herausgearbeitet und analysiert.

Die Brisanz liegt darin begründet, dass durch die steigende Flut von Daten, die heute teilweise den Terabyte-Bereich schon verlassen und den Petabyte-Bereich innerhalb eines Unternehmens erreicht haben, die Entscheidungsträger bei nicht funktionsfähigen oder zu langsamen Management-Information-Systemen (MIS) nicht in der Lage sind, die relevanten Informationen zu finden und zu nutzen. Durch die zunehmend kundennahen Produktionsstätten, bei denen Informationen detailliert für leistungsstarke, kundengerechte Fabrikation für den Standort und für den Mutterkonzern vorliegen müssen, sowie durch die wachsende Verbreitung von Daten unterschiedlichster Semantik über das Internet entwickelt sich die Notwendigkeit, die Performance der entscheidungsunterstützenden Systeme effizient zu steigern. Nur durch eine effektive Auswertung der vorhandenen relevanten Informationsbestände, können diese Daten für strategische und taktische Entscheidungen genutzt und damit ein Vorteil gegenüber der Konkurrenz erzielt werden.

Genau bei dieser Problematik, die strategischen Informationen zeit- und zielgenau aus dem global verteilten Unternehmenswissen kostengünstig zu finden und auszuwerten, kommen die Metadaten zum Tragen, da sich trotz einer immer schnelleren Datenautobahn nicht zwangsläufig eine Verbesserung der Integration und Verarbeitung der Daten und Informationen einstellt.

Die Schwierigkeit mangelhafte, unvollständige oder sogar falsche Verarbeitungsschritte und Interpretationen, ausgelöst durch willkürlich abgelegte Datensätze, zu vermeiden, wird durch das zentrale Element der Metainformation, der beschreibenden Information, gelöst. Gerade durch die Beschreibung der Daten können aus den Ressourcen des Internets, den unternehmensinternen operativen Datenbanken und Data Warehouses sowie den öffentlichen Verwaltungen Informationen katalogisiert werden, um durch das anschließende schnelle Wiederauffinden einen maximalen Nutzen zu erzielen.

Ein auftretendes Hindernis ist die unterschiedliche Metadatenstruktur, ausgelöst durch die individuellen Metamodelle der verschiedenen Data Warehouse-Tool-Anbieter und die damit verbundene, gravierende Inkompatibilität zwischen den Applikationen durch die fehlende Metadaten-Interoperabilität. Während der reine (Nutz-)