Association for Information Systems
AIS Electronic Library (AISeL)

Wirtschaftinformatik Proceedings 2005 | Wirtschaftinformatik

2-10-2005

Integration von Subsystemen in einem Active Data Warehouse

Carsten Felden
Universität Duisburg-Essen

Follow this and additional works at: http://aisel.aisnet.org/wi2005

Recommended Citation
Felden, Carsten, "Integration von Subsystemen in einem Active Data Warehouse" (2005). *Wirtschaftinformatik Proceedings 2005*. Paper 73.
http://aisel.aisnet.org/wi2005/73

This material is brought to you by the Wirtschaftinformatik at AIS Electronic Library (AISeL). It has been accepted for inclusion in Wirtschaftinformatik Proceedings 2005 by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

In: Ferstl, Otto K, u.a. (Hg) 2005. *Wirtschaftsinformatik 2005: eEconomy, eGovernment, eSociety*; 7. Internationale Tagung Wirtschaftsinformatik 2005. Heidelberg: Physica-Verlag

ISBN: 3-7908-1574-8

© Physica-Verlag Heidelberg 2005

Integration von Subsystemen in einem Active Data Warehouse

Carsten Felden
Universität Duisburg-Essen

Zusammenfassung: Die Informationsflut aus dem World Wide Web können Entscheidungsträger in Unternehmen schwer bewältigen. Das Marktdateninformationssystem verwendet einen Filter, der relevante Informationen für ein Unternehmen bestimmt und nur diese im Data Warehouse ablegt. Durch eine, in dieser Arbeit vorgestellte, Erweiterung eines bestehenden Marktdateninformationssystems werden Entscheidungsträger automatisch über wichtige Sachverhalte personalisiert informiert. Die unterschiedlichen Subsysteme der Architektur werden in unterschiedlichen Modellierungssprachen beschrieben und gepflegt. Zur Integration dieser Subsysteme wird eine Metamodellierung vorgeschlagen, welche die Konstruktion einer konsistenten Architektur für das Marktdateninformationssystem, um Wartung und Pflege im Systemlebenszyklus zu erleichtern, erläutert.

Schlüsselworte: Active Data Warehouse, Personalisierung, Metamodellierung

1 Einleitung

Die Informationsversorgung von Entscheidungsträgern ist mit Problemen behaftet, die zunächst aus einer Unterversorgung resultieren und in heutigen Zeiten durch eine Überversorgung gekennzeichnet sind [Kamp02, S. 9]. Durch das World Wide Web (WWW) werden unternehmensexterne Informationsquellen erschlossen, welche Textdokumente, im Sinne qualitativer Daten, als Träger der Informationen enthalten. Die umfangreichen qualitativen Daten können aus den Informationsressourcen nicht problemlos extrahiert werden [Hack99, S. 5].

Marktdaten haben entscheidenden Einfluss auf Unternehmen, wenn diese zeitnah genutzt werden. Durch geeignete Formen der Selektion und Repräsentation wichtiger Informationen werden Entscheidungsträger befähigt, ihre Aufgaben schneller zu erfüllen. Zu diesem Zweck kann ein Active Data Warehouse eingesetzt werden. Dieses versorgt Nutzer zeitnah mit entscheidungsrelevanten Informationen. Die verschiedenen Komponenten eines solchen Systems werden in Kapitel 3 erläutert. Jedoch erzeugt eine solche Erweiterung der Architektur Subsysteme innerhalb des Informationssystems. Keines der individuellen Modelle kann Datenobjekte des anderen Subsystems darstellen. Aus diesem Grunde zeigt Kapitel 4, wie mittels

Metamodellierung ein Inhalteraum konstruiert wird, der Subsysteme semantisch integriert und dadurch einen geschlossenen Ansatz ermöglicht.

2 Ausgangslage

Auf Grund der Liberalisierung der Energiemärkte innerhalb der Europäischen Union sah sich die Energiebranche neuen Herausforderungen gegenübergestellt. Die Beschaffungsoptimierung zur Kostenreduzierung umfasst nicht mehr nur die Beschaffung von Primärenergien, die in den eigenen Kraftwerken in elektrischen Strom umgewandelt werden, sondern auch die Beschaffung von Strom selbst. Die Möglichkeit des Handels mit Strom führte zu einer großen Vielfalt an neuen Produkten und potenziellen Handelspartnern. Viele Unternehmen trugen dieser Entwicklung Rechnung, indem sie einen sogenannten Trading Floor einrichteten, dessen Mitarbeiter sich vorwiegend der Aufgabe des Energiehandels widmen. Zu ihrer Unterstützung wurde ein Marktdateninformationssystem (MAIS) bereitgestellt, das allen Mitarbeitern des Handelsbereichs als Datenbank für Marktdaten zur Verfügung steht. Für die anfallenden Entscheidungsaufgaben können auf Grundlage der Marktdaten Marktanalysen, statistische Auswertungen und Preisprognosen betrachtet werden. Dies erfordert die Integration heterogener Datenformate aus internen oder externen Quellen in einer Datenbank.

Jedoch ist zu bedenken, dass ein solches System eine Informationsüberversorgung der Entscheidungsträger forciert. Marktanalysten benötigen zwar eine Vielzahl an Daten, jedoch müssen sie diese im ursprünglichen System eigenständig selektieren. Das bedingt einen entsprechenden Zeitaufwand, dem im oftmals hektischen Handel nicht immer entsprochen werden kann. Da die Entwicklung des MAIS in mehreren Stufen erfolgt, an deren Ende jeweils ein funktionsfähiges Informationssystem steht, können Erkenntnisse über die Auswirkungen des Systemeinsatzes als Ausgangspunkt für Weiterentwicklungen betrachtet werden. Sie konzentriert sich darauf, quantitative und qualitative Daten nicht nur zu integrieren, sondern diese auch bedarfsgerecht Entscheidungsträgern zur Verfügung zu stellen. Dazu ist eine Anwenderprofilierung durchzuführen. Das ermöglicht nicht nur eine frühzeitige Information über Marktereignisse, sondern darüber hinaus auch über frühe Signale des Marktes.

Um beide Anforderungen befriedigen zu können, wird die Architektur um zwei Bausteine erweitert. Der erste Baustein besteht aus Benutzerprofilen, um personalisiert Informationen zur Verfügung zu stellen zu können. Der zweite Baustein besteht aus Frühindikatorenprofilen, um diese in Texten identifizieren zu können. Dabei handelt es sich um zwei grundsätzlich unterschiedliche Konzepte. Benutzerprofile ermöglichen eine Zeitersparnis, da sie dem Analysten Informationen zielgerichtet zur Verfügung stellen und dadurch der Informationsüberversorgung entgegenwirken. Frühindikatorenprofile ermöglichen einen Zeitvorteil, da An-

wender, durchaus profilorientiert, frühzeitig über potentielle Entwicklungen informiert sind. Durch den Wandel von einer Pull- zu einer Push-Architektur entwickelt sich das MAIS zu einem Active Data Warehouse.

3 Gestaltung einer Push-Architektur eines Marktdateninformationssystems

Im folgenden Kapitel wird zunächst das Basissystem vorgestellt. Dies wird um Module zur personalisierten Informationsversorgung sowie zur Identifikation von Frühindikatoren in qualitativen Textdokumenten erweitert.

3.1 Metadaten basierte Integration externer Daten und Erweiterung zum Active Data Warehouse

Die Integration und Repräsentation interner quantitativer und qualitativer Datenbestände externer Herkunft kann durch ein MAIS vorgenommen werden. Das System wurde in einem Projekt der VEW Energie AG und der Universität Duisburg realisiert. In diesem Zusammenhang werden Extraktoren, Filtertechnik, Datenbank sowie die Repräsentation der Daten als Ganzes betrachtet und eine weitgehend automatisierte und fachspezifische Informationsversorgung vorgenommen [Feld02, S. 57f; Feld02, S. 142f]. Die interdependenten Beziehungen der Systemkomponenten des MAIS sind in Abbildung 1 dargestellt.

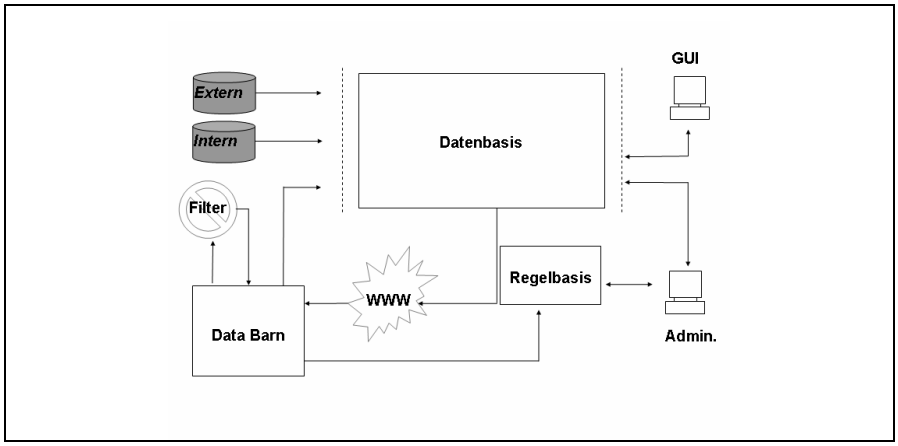

Abbildung 1: Aufbau des Marktdateninformationssystems [Feld02, 57]

Nach einer metadatenbasierten Suche werden alle potenziell relevanten Dokumente aus dem WWW in der Data Barn klassifiziert [Feld02, S. 58]. In der Daten-

basis werden ausschließlich die relevanten Dokumente abgelegt [Feld02, S. 128f]. Die Prüfung geschieht im Rahmen eines Filterprozesses, indem ein Künstliches Neuronales Netzwerk die Unterscheidung von relevanten und nicht relevanten Dokumenten vornimmt. Dieser Filter beinhaltet lang- und mittelfristig stabile Kriterien und leitet nur spezifische Informationen in die Datenbasis. Die Filterkriterien sind in der Regelbasis abgelegt [Feld02, S. 58]. Grenzen sind der Automatisierung der Datenintegration dahingehend gesetzt, als das eine manuelle Prüfung der Filterklassifikation durchgeführt werden muss. Diese kann nicht hundertprozentig garantiert werden. Die Repräsentation der relevanten Dokumente erfolgt über das Graphical User Interface (GUI) in Form eines Starfield Display [LeRi99], bestehend aus Symbolen in einem zweidimensionalen Koordinatensystem [Feld02, S. 76f]. Über die Abszisse wird die Zeit und über die Ordinate der Relevanzwert wiedergegeben. Eine Einordnung erfolgt somit auf der Grundlage von Erscheinungsdatum[1] und Relevanzwert, welcher mit dem Künstlichen Neuronalen Netz (KNN) bestimmt wurde.

Obwohl eine Selektion der Textdokumente durch das KNN vorgenommen wird, können aus Gründen der Zeitknappheit nicht alle Dokumente von den Entscheidungsträgern gesichtet werden. Somit rückt die individuelle Versorgung der Entscheidungsträger mit interessanten Informationen ins Zentrum der weiteren Betrachtung. Insbesondere durch eine reaktive Versorgung mit aktuellen Informationen werden bei zeitkritischen Entscheidungen Wettbewerbsvorteile erschlossen. Diese Transaktionen werden automatisiert durch das Active Data Warehouse durchgeführt [Frit02, S. 5]. Dabei wird die reaktive Komponente als Erweiterung des MAIS verstanden. Ein Active Data Warehouse überwacht bestimmte Sachverhalte und führt durch Trigger bestimmte Aktionen aus [Schl00, S. 92]. Im Folgenden wird das Active Data Warehouse genutzt, um Nachrichten an Nutzer zu senden, wobei die zu überwachenden Daten in quantitative und qualitative Daten unterschieden werden. Zunächst wird eine Lösung für quantitative Daten aufgezeigt, um im Anschluss Möglichkeiten für qualitative Daten zu erläutern.

Für *quantitative Daten* führen Trigger-Mechanismen Aktionen auf Basis eines Regelwerkes aus. Trigger bestehen aus einer formulierten Situation und einer Datenbankprozedur, die bei Eintritt der spezifizierten Situation abgearbeitet wird [Schl00, S. 92]. Hierdurch können Geschäftsregeln in ein Informationssystem integriert werden [BeKS02, S. 244, S. 246]. Die Realisierung geschieht durch Event-Condition-Action (ECA)-Regeln [Schl00, S. 3, S. 100; Frit02, S. 5ff]. Die Ereigniskomponente (Event) enthält vordefinierte Ereignisse, welche den Regeldurchlauf auslösen. Die Bedingungskomponente (Condition) enthält die zu prü-

[1] Das Erscheinungsdatum ist das Datum, an dem ein Dokument auf einem Server publiziert wurde [Feld02, S. 77]. Dieses wird bei Erfassung auf Greenwich-Zeit standardisiert.

fenden Kriterien. Sind die Ereignis- und Bedingungskomponenten erfüllt, wird die in der Aktionskomponente (Action) enthaltene Reaktion ausgeführt.

Schwieriger gestaltet sich der Sachverhalt bei *qualitativen Daten*. Nach Schätzungen liegen bis zu 80 Prozent der betrieblichen Informationen in Textdokumenten vor [Sull01, S. 56]. Sind diese zudem externer Herkunft, kommt ihnen eine besondere Bedeutung zu, da sie wichtige und aktuelle Informationen beinhalten können. Durch den beschriebenen Integrationsprozess werden ständig relevante Textdokumente aus dem WWW in das Data Warehouse importiert. Insbesondere zeitkritische Entscheidungen werden durch eine intensive Datensuche verzögert. Das hat Auswirkungen auf die Entscheidungsträger, welche die Aufgabe der Dokumentensichtung in ihrem begrenzten Zeitportfolio als bedrohlich empfinden [Meie03, S. 475].

Im Sinne eines Active Data Warehouse werden neue Informationen direkt an bestimmte Nutzer weitergeleitet [Ecke03, S. 8]. Bei diesem Push-Dienst sind die rollenspezifischen und individuellen Informationsbedürfnisse der einzelnen Nutzer zu berücksichtigen, damit der Informationsflut entgegen gewirkt werden kann. Da die Textdokumente meist unstrukturiert sind, ist der Interessantheitsgrad schwer zu bestimmen. ECA-Regeln sind auf klar definierte Kriterien ausgelegt, welche für Textdokumente zunächst nicht vorliegen. Das Text Mining unterstützt die Anwendung von ECA-Regeln. Aus den einzelnen Dokumenten werden Metadaten extrahiert und für weitere Analysen genutzt [VoGu01, S. 254ff]. Diese Metadaten bilden die Grundlage, um eine nutzerspezifische Kennzahl zu berechnen. Diese repräsentiert den individuellen Interessantheitsgrad eines Dokumentes. ECA-Regeln überprüfen die Kennzahl und führen gegebenenfalls Aktionen aus.

3.2 Ontologiebasierte Profilentwicklung

Das Ziel dieses Verfahrens ist eine inhaltsbezogene und weitgehend automatische binäre Textklassifikation in für den Nutzer interessante und uninteressante Dokumente. Da derzeit zu den meisten Textdokumenten aus dem WWW keine Metadaten über den Inhalt verfügbar sind, werden die Terme eines Dokumentes zur inhaltlichen Erschließung verwendet. Über die Technik der Indizierung kann der Inhalt beschrieben werden. Die durch das MAIS bereits vorgenommene Indizierung wird leicht modifiziert verwendet.

Der durch das MAIS ausgelöste Informationsstrom an Textdokumenten wird mit einem individuellen Informationsbedarf verglichen und weitergeleitet, wenn dieser befriedigt werden kann [FoDu92, S. 52f]. Der Informationsbedarf kann auf Grund relativ stabiler und spezifischer Aufgabenbereiche von Mitarbeitern bestimmt werden. Dabei werden Dokumente nur einmalig nach dem Import in das Data Warehouse überprüft [AbKW02, S. 89]. Der Informationsbedarf wird somit nur mit aktuellen Informationen verglichen und eine Filterung auf verschiedenen Stufen durchgeführt. Die Stufen können in einen Situationsfilter und einen Benutzerfilter

unterteilt werden [Meie03, S. 485ff]. Der Situationsfilter zeichnet sich vor allem durch die Situierung eines Unternehmens aus und wird durch das KNN des MAIS bereits realisiert. Durch den Benutzerfilter wird eine inhaltlich orientierte Analyse ermöglicht, in der die Informationen in einem Dokument mit dem Informationsbedarf eines Nutzerprofils verglichen werden. Der Benutzerfilter basiert auf Nutzerprofilen, welche bezüglich der Rolle und der Personalisierung unterschieden werden [Meie03, S. 486]. Die Rolle beinhaltet organisatorische und soziale Aspekte, die sich objektiv als komplexes Gebilde von Erwartungen an den Inhaber einer Position manifestiert.

Der Fokus liegt im Folgenden auf der Entwicklung einer Schnittstelle zwischen den Dokumenten und den Nutzern. Diese ermöglicht den Nutzern die selbstständige Erstellung eines Profils. Die Dokumente und Nutzerprofile werden hierbei durch Vektoren repräsentiert und miteinander verglichen. Eine Erstellung des Nutzerprofils über die Angabe jedes einzelnen Deskriptors ist aufwändig. Synonym verwendete Deskriptoren sind zudem explizit zu bestimmen. Da Deskriptoren mehrere Bedeutungen haben können, liegen insgesamt die Ergebnisse der Bewertungsmaßgaben *Precision* und *Recall* auf niedrigem Niveau [Sull01, S. 472]. Deshalb erfolgt der Zugriff auf die Dokumente über eine Ontologie, die Konzepte enthält. Um Konzepte zu bestimmen, sind Synonyme und Verbundereignisse automatisch zu identifizieren. Zu beachten ist, dass die Konzepte neben den Synonymbeziehungen auch hierarchische oder andere Beziehungen zwischen Deskriptoren enthalten können. Obwohl dies automatisch durchgeführt werden kann, ist eine manuelle Überprüfung vorzunehmen [AbKW02, S. 132]. Die Interdependenzen des gesamten Klassifikationsverfahrens sind in Abbildung 2 dargestellt.

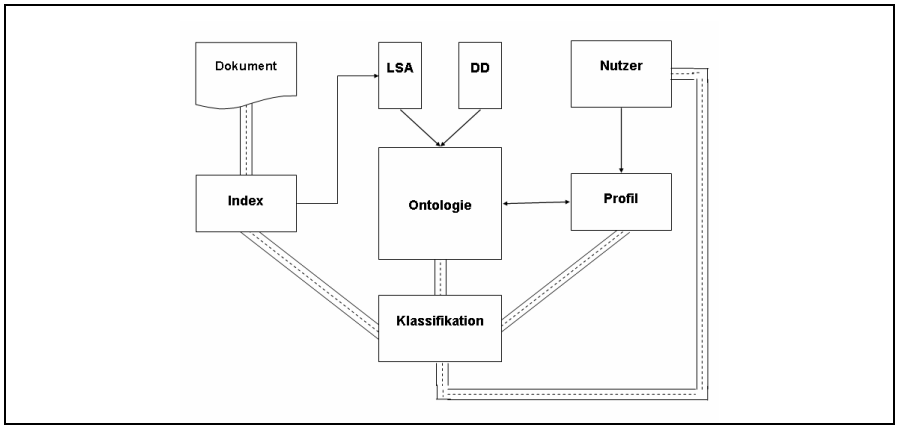

Abbildung 2: Aufbau des Klassifikationsverfahrens

Die gestrichelten Kanäle symbolisieren hierbei die ständig aktiven Verbindungen des eigentlichen Verfahrens. Die Pfeilverbindungen symbolisieren vor allem einmalige Abläufe, die nur unter bestimmten Umständen wiederholt werden. Der

Prozess der Ontologieerstellung wird einmalig vollzogen und nur in bestimmten Abständen oder nach besonderen Ereignissen erneut durchlaufen, da die Erstellung mit hohem Aufwand verbunden ist. Durch die zusätzliche Hinterlegung von unternehmensspezifischem Fachwissen [Feld02a, S. 750], in Form der unternehmensspezifischen Terminologie [AbKW02, S. 131], kann die Ergebnisqualität des zu entwickelnden Klassifikationsverfahrens verbessert werden. In diesem Prozess werden die Zusammenhänge und Bezeichnungen des Data Dictionary berücksichtigt, wodurch eine einheitliche Terminologie geschaffen wird und die Verknüpfung mit den quantitativen Daten im Data Warehouse realisiert ist [Sull01, S. 155ff]. Dies führt dazu, dass Ontologie, Data Dictionary und Glossar verschmelzen müssen. Auch hier besteht die Einschränkung der Automatisierung dahingehend, dass die Ontologieentwicklung mit ihren Verknüpfungen manuell verifiziert werden muss.

Diese Ontologie ist Basis für ein Benutzerprofil. Sie beinhaltet semantisches Wissen, welches maschinell verarbeitbar ist und somit die Kommunikation zwischen Mensch und Maschine verbessern kann [Fens98, S. 11]. Die Konzepte einer Ontologie bilden ein abstraktes Modell. Dieses Modell gilt nur für einen abgegrenzten Bereich, da ansonsten die Beziehungen zwischen den Konzepten nicht mehr gelten. Somit werden qualitativ hochwertige Ergebnisse nur erzielt, wenn die Textdokumente dem speziellen Themenbereich zugehören, für welche die Ontologie erstellt wurde. Die in einer Ontologie enthaltenen Konzepte werden durch Beziehungen miteinander verbunden [ArCT03, S. 62; KaSh00, S. 12]. Durch die Strukturierung der Konzepte entsteht eine unternehmensspezifische Taxonomie [Feld02a, S. 754]. Die Konzepte bestehen aus einer Konzeptbezeichnung, einer kurzen Konzeptbeschreibung [AbKW02, S. 132], einer dem Konzept zugeordneten Menge synonym verwendeter Begriffe und einer Menge direkt untergeordneter Konzepte. Die erstellten Konzepte können manuell in die zwei Mengen der Synonyme und der untergeordneten Konzepte eingeteilt werden. Über die Hierarchie wird anschließend der benötigte Detaillierungsgrad zur Beschreibung der Dokumente bestimmt.

Die Beschreibungen aus dem Data Dictionary werden jedoch nicht vollständig für die Ontologie benötigt. Quantitative Beziehungen werden nicht berücksichtigt. Zudem werden die bereits gebildeten Thesauren zur Modifikation der Indizes für das KNN berücksichtigt. Das Nutzerprofil besteht aus Themenbereichen und Regeln. Beide entnehmen die Terme aus den Konzepten der Ontologie, werden jedoch unterschiedlich überprüft. Ein Themenbereich enthält mehrere Konzepte, und der hieraus entstehende Vektor wird mit dem Index der Dokumente verglichen. Es besteht keine Notwendigkeit, dass alle Konzepte in einem Index existieren, damit ein Dokument als interessant klassifiziert wird. Eine nutzerspezifische Regel enthält ebenfalls Konzepte, um auch alternative Begriffe identifizieren zu können. Alle Konzepte müssen in einem Index vorkommen, damit ein Dokument als interessant klassifiziert wird [LiHs02, S. 467f]. Bevor die Nutzer die Personali-

sierung vornehmen, werden ihnen bestimmte Themenbereiche bezüglich ihrer Rolle im Unternehmen zugewiesen [AbKW02, S. 157].

In der Dokumentenanalyse werden die neu in das Data Warehouse importierten Dokumente auf den individuellen Interessantheitsgrad geprüft. Unter Nutzung der Ontologie werden die Dokumente über die Konzepte abgebildet und mit den Nutzerprofilen verglichen. Ein neues Dokument wird zunächst auf die einzelnen Themenbereiche und im Anschluss daran auf die Regeln für jedes Nutzerprofil überprüft. Jeder Vektor eines Themenbereiches wird mit dem Index eines Dokumentes verglichen und hierzu der Kosinus-Koeffizient als Ähnlichkeitsmaß gebildet [Sull01, S. 335]. Die in den Regeln enthaltenen Konzepte werden mit dem Index des Dokumentes verglichen, wobei diese Konzepte vollständig oder gar nicht enthalten sein müssen. Über den Kosinus-Koeffizienten wird somit die Ähnlichkeit eines Index mit den einzelnen Vektoren der Themenbereiche ermittelt [FoKL98, S. 5]. Jedoch berücksichtigt der Koeffizient nicht die Gewichtung der einzelnen Konzepte, denn er prüft nur die Existenz eines Konzeptes in einem Dokument [Ferb03, S. 74]. Da die im Data Warehouse befindlichen Dokumente durch das KNN auf eine Mindestrelevanz für das Unternehmen geprüft worden sind, ist der Kosinus-Koeffizient ausreichend, um thematische Ähnlichkeiten zu bestimmen. Nach Abschluss der Teilanalyse liegen zu jedem Themenbereich eines Profils die entsprechend ermittelten Ähnlichkeitswerte vor. Somit verbleiben die Regeln, welche anschließend zu prüfen sind. Regeln enthalten weniger Konzepte als die Themenbereiche, wobei die räumliche Nähe der Konzepte zueinander eine wichtige Rolle einnehmen kann [Pull01, S. 8]. Diese wird durch die bisher verwendeten Indizes der Dokumente nicht berücksichtigt [Pull01, S. 3]. Bei der Indizierung gehen jegliche Informationen über die Anordnung der Deskriptoren verloren. Aus diesem Grund werden die Indizes nicht zur Prüfung der Regeln verwendet.

Bei der Überprüfung einer Regel werden drei aufeinander folgende Sätze berücksichtigt. Diese Bedingung ist weniger restriktiv. Sind die Sätze relativ kurz, besteht die Möglichkeit, dass die gesuchten Konzepte in einem nachfolgenden Satz enthalten sind. Beginnend mit den ersten drei Sätzen eines Dokumentes wird überprüft, ob die Konzepte in den Sätzen enthalten sind. Hierbei werden die drei Sätze temporär als Ganzes betrachtet und die Konzepte in diesen ermittelt. Die Reihenfolge der Konzepte in den Sätzen ist unwichtig. Nach der Prüfung wird der erste Satz entfernt und für diesen der vierte Satz in die Prüfung einbezogen. Dann werden diese drei Sätze als Ganzes betrachtet und untersucht. Somit wird die Prüfung Schritt für Schritt bis zum letzten Satz des Dokumentes durchgeführt.

Nach vollständiger Dokumentanalyse ist festzustellen, ob der Nutzer durch die zuvor im Nutzerprofil spezifizierte Nachricht zu informieren ist. Die Auslösung der Benachrichtigung wird durch einen Trigger vorgenommen. Dieses ist möglich, da im Rahmen der Themen- und Regelüberprüfung die Ergebnisse für jedes Dokument in Form von Werten abgelegt wurden. Für jeden Themenbereich liegt ein Kosinuswert als Betrag aus dem Intervall [0, 1] vor [GeHK01, S. 41]. Da die Re-

geln nur auf vollständige Erfüllung überprüft werden, liegen entweder die Werte Null oder Eins vor. Ein Trigger wird für jedes Profil abgeleitet und nach vollständiger Analyse eines Dokumentes (Event) ausgeführt. Dieser Trigger enthält die Schwellenwerte aus den Nutzerprofilen. Ist mindestens ein Schwellenwert überschritten oder eine Regel erfüllt (Condition), wird eine Benutzernachricht mit den entsprechenden Informationen versendet (Action). Somit kann ein Active Data Warehouse auf Basis qualitativer Daten realisiert werden. Hierbei ist es von Interesse, vom Analysten eine Rückkopplung durch eine Bewertung der zur Verfügung gestellten Textdokumente zu erhalten. Dadurch kann die Funktionsweise des Triggers optimiert werden.

3.3 Profilbasierte Identifikation von Frühindikatoren

Sowohl quantitative als auch qualitative Daten im Data Warehouse sind gemäß des Integrationsprozesses vergangenheitsorientiert. Sie repräsentieren entsprechend Ihrer Aufgabe im MAIS betriebswirtschaftliche Ereignisse im Zeitverlauf. Die Analyse von Textdokumenten zur Identifikation von Frühindikatoren wendet die Blickrichtung in die Zukunft. Frühindikatoren geben Hinweise auf potenzielle wirtschaftliche Entwicklungen [Anso80]. Um den Analysten einen Zeitvorsprung zu ermöglichen, wird ein weiteres Modul in der Architektur des MAIS angelegt, um potenzielle Auswirkungen der identifizierten Frühindikatoren aufzuzeigen. Zusätzlich ermöglicht es auch eine thematische Kopplung einzelner Frühindikatorenmodelle an entsprechende Berichtskomplexe im Data Warehouse. Das unterstützt die Bindung an entsprechende Benutzerprofile, um Ergebnisse anwenderorientiert zur Verfügung zu stellen.

Eine gleiche zukunftsorientierte Blickrichtung liegt auch dem Konzept der Simulation zu Grunde. Es wird nicht nur in der Betriebswirtschaft eingesetzt, um beispielsweise Planwerte als Zielvorgaben unternehmerischer Tätigkeit zu bestimmen oder die Auswirkung unterschiedlicher Einflussfaktoren auf das Marktgeschehen zu bewerten. Jedoch sind bei der Anwendung einer Simulationstechnik Grenzen gesetzt. Ein allgemein gültiges, unternehmensweites Simulationsmodell lässt sich praxisgerecht nicht erstellen. Die Dynamik der Umwelt sowie der Umfang der Modellparameter sind bei einer umfassenden Modellierung zu intensiv. Des Weiteren ist zu beachten, dass ein Simulationsmodell nicht nur der Dynamik des Modells entsprechen soll, sondern auch der Dynamik der Modellumwelt unterliegt. Dies erfordert einen entsprechenden Pflege- und Wartungsaufwand. Um diesen Aufwand einzuschränken, sind Simulationsparameter derart zu gestalten, dass die Dynamik der Umwelt keinen dominierenden Einfluss erhält und somit die Systemkomplexität unnötig zunimmt. Das konkret im MAIS integrierte Modul basiert auf der Simulationsmethode *System Dynamics*. Die Auswahl ist damit begründet, dass System Dynamics bei der Modellierung zusätzlich zu quantitativen Einflussgrößen auch qualitative explizit erfasst [Ster00; Forr69; Mead72]. Auf Grund der oben

genannten Einschränkungen, wird ein Feedback-Diagramm gemäß System Dynamics eingesetzt.

Zunächst ist ein Modellbereich abzugrenzen. Eine thematische Verbindung zu einem Berichtskomplex grenzt den Modellbereich ein und macht die Parameterkomplexität beherrschbar. Diese orientiert sich analog zur bisherigen Vorgehensweise an den Metadaten und dem Glossar im Kontext zur Ontologie des multidimensionalen Datenmodells. Dadurch ist zum einen Begriffseindeutigkeit gewährleistet, zum anderen erfolgt eine Kopplung der Frühindikatorenprofile an die jeweiligen Berichtskomplexe. Bezieht sich ein Bericht zum Beispiel auf die Nachfrage nach Energiederivaten, ist dies als zentrale Information im Modell zu erfassen. Durch eine Informationsbedarfsanalyse sind Variablen und Konstanten zu identifizieren, die einen Einfluss auf die beispielhaft aufgeführte Nachfrage haben. Das Ergebniss der Informationsbedarfsanalyse wird in einem Feedback-Diagramm beschrieben. Die Identifikation von Wirkungszusammenhängen in qualitativen Daten setzt voraus, dass die Textsemantik automatisiert erkennbar ist. Auch hier besteht das Problem, dass identische Aussagen durch variierende Begriffe beschrieben werden können. Das erfordert ein Verfahren, Inhalte trotz variierender Ausdrucksweisen zu erkennen und Wirkungszusammenhänge zu bestimmen. Die Funktionsweise wird in Abbildung 3 gezeigt.

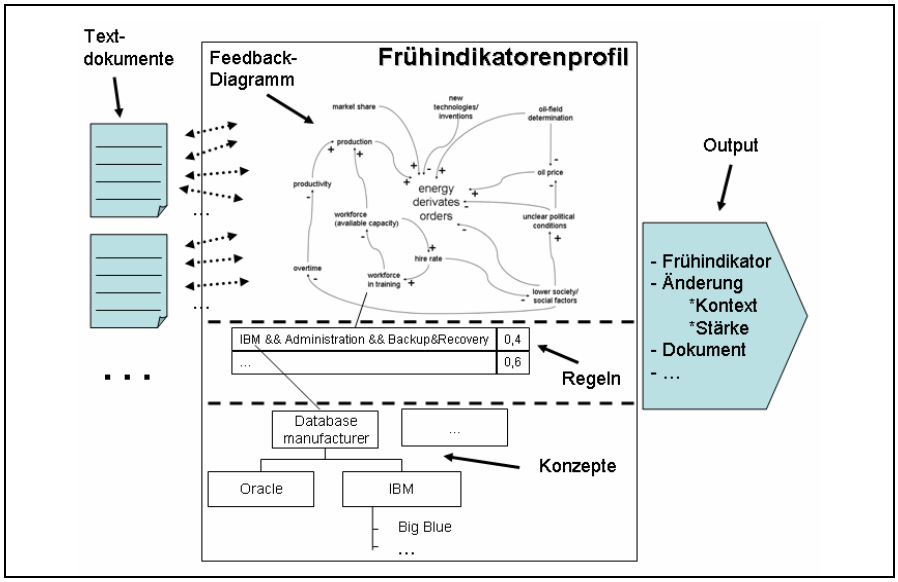

Abbildung 3: Funktionsweise des Frühindikatorenprofils

Das Architekturmodul *Frühindikatorenprofile* erfasst disjunkte Feedback-Diagramme zu einem Themenkomplex mit entsprechend generierten Regelbibliotheken der Parameter, die auf Konzepten der Ontologie basieren. Dort sind zusätz-

lich die Einflussausprägungen zwischen diesen dokumentiert. Wird ein neues Textdokument in der Datenbank erfasst, ist anhand der Deskriptoren zu prüfen, ob der Wert einer Modellvariablen zu verändern ist. Dem Gedanken der Regelprüfung des Benutzerprofils folgend wird die Prüfung direkt im Dokument ausgeführt, nicht auf dem erzeugten Index. Abbildung 3 zeigt beispielhaft die Regelprüfung, ob Mitarbeiter zu einer Datenbankschulung geschickt worden sind. Ist ein Suchterm gefunden, ist der Variablenwert im Modell um die zugeordnete Quantifizierung zu verändern und der Anwenderkreis zu informieren, in dessen Profil der Bericht als Merkmal erfasst ist. Dieser Output umfasst z. B. den identifizierten Frühindikator, die ausgeführte Änderung sowie das Dokument, in dem der Indikator identifiziert wurde.

Die zuvor beschriebene Erweiterung der Architektur erzeugt Subsysteme innerhalb des MAIS. Diese werden durch Beschreibungssprachen mit einer unterschiedlichen Semantik dargestellt. Um einen konsistenten Ansatz zu entwickeln, sind diese Subsysteme miteinander zu verknüpfen.

4 Integration der Architektur-Subsysteme

Grundsätzlich wird eine solche Verknüpfung bereits seit einiger Zeit unter der Bezeichnung *Unternehmensmodell* diskutiert [Fran99; JJPS96]. Dabei geht es darum, verschiedene sprachliche Systeme, die naturgemäß innerhalb eines Unternehmens bestehen, zu integrieren. Zu diesem Zweck werden, im weiteren Verlauf des Beitrages orientiert an [BeBC+03, S. 27ff], Inhalteräume gestaltet. Hier ist die semantische Vorgabe eine Cube-Struktur des Data Warehouse. Darum wird ein Raum gebildet, in dem Benutzerprofile und Frühindikatorenprofile entsprechend objektiv zugeordnet werden, da sie sich immer auf den Inhalt einer Cube-Struktur beziehen. Eine solche Vorgehensweise integriert die verschiedenen sprachlichen Systeme semantisch durch eine Vielzahl von disjunkten Inhalteräumen im MAIS.

Zunächst wird eine Sprache semi-formal mittels Metamodellierung, welche in Abschnitt 3.1 vorgestellt wird, beschrieben. Zur Spezifikation wird dann die im Subsystem angewendete Sprache in das Metamodell eingesetzt. Das Metamodell selbst basiert auf einem Meta-Meta-Modell, das die Beschreibungselemente des Modells definiert. Dies ist die Ebene, auf der die unterschiedlichen Modellierungssprachen zusammengeführt werden können, da eine einheitliche Spezifikation für alle Modelle zu Grunde gelegt wird. Als oberste Integrationsebene wird im Folgenden die *Meta-Object-Facility* der *Object Management Group* (OMG) genutzt und in Abschnitt 3.2 beschrieben. Abschließend schildert Abschnitt 3.3 die Integration der unterschiedlichen Sprachen auf Basis des einheitlichen Meta-Meta-Modells und der darin enthaltenen Konzepte. Dabei handelt es sich im Ergebnis nicht um ein Unternehmensmodell. Jedoch folgt dieses Konzept dem gleichen

Prinzip und unterstützt eine konsistente Architektur für ein Analytisches Informationssystem.

4.1 Meta-Modellierung

Im Rahmen einer Modellierung werden die Inhalte einer Domäne mit den Konzepten, die durch eine Modellierungssprache zur Verfügung gestellt werden, beschrieben. Die Metamodellierung ist eine Ebene höher angesiedelt. Sie erlaubt die Modellierung der Modellierungssprache. Dabei ist die Sprache auf bestimmte Problembereiche zugeschnitten. Dies drückt sich in der jeweiligen Modellmetapher aus. Eine solche Metapher symbolisiert den Entwicklungszweck der Sprache [OMG02, S. 1-1ff].

Um eine Modellierungssprache spezifizieren zu können, muss ein Metamodell generiert werden. Dazu wird eine weitere Sprache benötigt: die Meta-Meta-Modellsprache. Dies bedeutet auch, dass unterschiedliche Modellebenen in die Betrachtung involviert sind. Im Schrifttum werden (allgemein akzeptiert) vier Ebenen unterschieden [OMG02, S. 2-3]. Die *Instanzebene* enthält die Tupel der Datenbank, die *Schemaebene* die Struktur des zu Grunde liegenden Datenmodells. Dabei handelt es sich in diesem Kontext um ein Star-Schema für die Datenbank, eine boolesche Liste für das Benutzerprofil und ein Feedbackdiagramm zur Identifikation der Frühindikatoren. Die *Metaebene* enthält die Beschreibungselemente, die auf der Schemaebene zu nutzen sind. Jedoch kann ein Metamodell nicht um die Bestandteile eines anderen Metamodells erweitert werden, um die jeweiligen deskriptiven Fähigkeiten zu erreichen. Dies erzeugt einen Widerspruch in der Modellmetapher. Keine dieser Modellierungsmethodiken wurde für andere Darstellungsinhalte geschaffen. Die Integration kann nur auf einer *Meta-Meta-Ebene* stattfinden, da dort die jeweiligen Metamodelle definiert werden.

Auch die OMG [OMG02, S. 2-3] orientiert sich an dieser vierfachen Unterteilung und hat z. B. mit dem Common-Warehouse-Metamodel (CWM) ein Referenzmodell für die Modellierung von Data-Warehouse-Systemen auf der Meta-Ebene geschaffen [OMG01]. Dieses basiert auf der Meta-Object-Facility (MOF), die für das weitere Vorgehen zu Grunde gelegt wird. Ihre Aufgabe ist die eindeutige formale Beschreibung des Meta-Modells. Dabei ist die MOF nicht nur Basis des CWM, sondern sämtlicher Methoden zur spezifizierenden Datenmodellierung [OMG02].

4.2 Meta-Object-Facility

Die Meta-Object-Facility ist eine objektorientierte Modellierungsmethode. Sie basiert auf der zuvor beschriebenen Vier-Schichten-Metamodell-Architektur [OMG02, S. 2-2f]. Ihre Aufgabe ist die Spezifikation von Schnittstellen in Anlehnung an die Common Object Request Broker Architecture (CORBA), um inter-

operierende Metamodelle zu definieren und zu manipulieren [OMG02, S. xvi]. Dazu wird die visuelle Syntax der Unified Modelling Language (UML) genutzt [OMG02, xii]. Sämtliche Spezifikationen der OMG sind Instanzen dieser MOF. Die geforderte Interoperabilität soll sämtliche Phasen der Applikationsentwicklung mit deren jeweiligen Metamodellen unterstützen. Dies bedeutet, dass solche Metamodelle eine Instanz der MOF bilden. Das kann in dieser Vollständigkeit nicht strikt eingehalten werden. Strikte Metamodellierung fordert, dass jedes Element einer nachrangigen Stufe Instanz eines Elementes der darüber liegenden Stufe ist. Die OMG hat in diesem Kontext erkannt, dass eine Instanziierung aller MOF-Elemente im Metamodell nicht möglich ist. Somit kann die MOF für die Definition von Metamodellen genutzt werden, die eine andere Philosophie bzw. andere Details aufweisen. Abbildung 4 zeigt einen Ausschnitt des MOF-Klassendiagramms, welches die Struktur und Semantik der Meta-Meta-Daten darstellt.

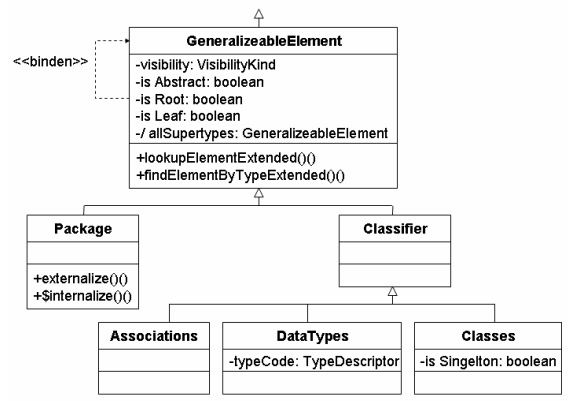

Abbildung 4: Ausschnitt der Meta Object Facility [OMG02, S. 3-12 (modifiziert)]

Diese Abbildung zeigt die Konstrukte der Metamodellierung. Die vier zentralen Modellierungskonzepte sind [OMG02, S. 2-6ff]:

- Klassen (Classes), welche die MOF-Meta-Objekte modellieren. Diese enthalten Attribute und Operationen.
- Assoziationen (Associations), welche binäre Beziehungen zwischen Meta-Objekten modellieren. Unter dem Aspekt, dass die Beziehung ggf. ein hierarchisches Verhalten aufweist, kann es sich bei der Assoziation auch um eine Aggregation handeln.
- Datentypen (DataTypes), die z. B. primitive Typen (Integer, String etc.) oder externe Typen (für nicht MOF-Spezifikationen erforderlich) definieren.
- Pakete (Packages), die das Modell in Module aufteilen.

Das MOF-Modell beschreibt die vom Meta-Modellierer nutzbaren Metadaten, um Klassen mit deren Attributen und Assoziationen als Relationen zwischen Klassen abbilden zu können. Diese Konstrukte sind ausreichend, um das Metamodell als abstrakte Syntax zu erstellen. Dazu ist es sinnvoll, die Syntax durch entsprechende Konsistenzbedingungen zu erweitern [OMG, S. 2-17ff]. Abgeleitet aus den Objekten *Class* und *Association* werden im Folgenden die Metamodelle für die ADAPT-Modellierung des Data Warehouse, die Benutzerprofilierung und System Dynamics eingeführt.

4.3 Meta-Modelle des Marktdateninformationssystems

Bei dem mit der MOF verknüpften Metamodell handelt es sich um die Basis für ein konzeptuelles beziehungsweise semantisches Modell. Die Erstellung eines solchen Modells ist eine systemunabhängige Tätigkeit [ElNa02], so dass die Basisanforderungen wertfrei auf der Metaebene modelliert werden können. Die Abbildung 5 zeigt das Metamodell der Modellierungssprache *Application Design for Analytical Processing Technologies* (ADAPT).

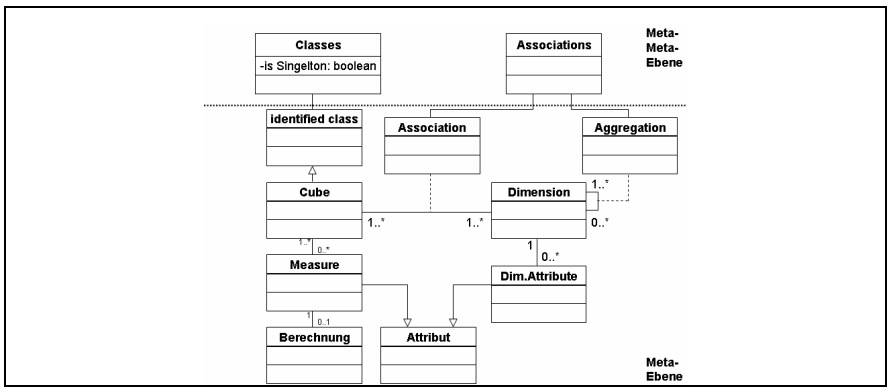

Abbildung 5: ADAPT-Meta-Modell [PrPe01, S. 78 (modifiziert)]

Das Metamodell beinhaltet die Kernelemente Würfel (Cube), Kennzahl (Measure), Hierarchieebene (Dimension) und Dimensionsattribute (Dim.Attribute). Der Würfel besitzt als Untersuchungsobjekt null bis mehrere Kennzahlen, welche die zu analysierenden Fakten repräsentieren. Die Hierarchieebenen sind in Halbordnungen angeordnet, wobei ein Würfel jeweils durch eine Menge von Basishierarchieebenen aufgespannt wird.

Die Basis für eine Personalisierung bildet immer ein sogenanntes Benutzerprofil [Runt00; Mert97]. Darin werden die gesammelten personenbezogenen Daten gespeichert. Mit Hilfe dieser Daten lässt sich der gesamte Ablauf von Transaktionen individualisieren. Um die abstrakten und persönlichen Daten erfassen zu können, muss ein wertfreies Metamodell bestehen, um ein systemunabhängiges Modell zu

erstellen. Abbildung 6 zeigt das Metamodell zu Gunsten einer Benutzermodellierung.

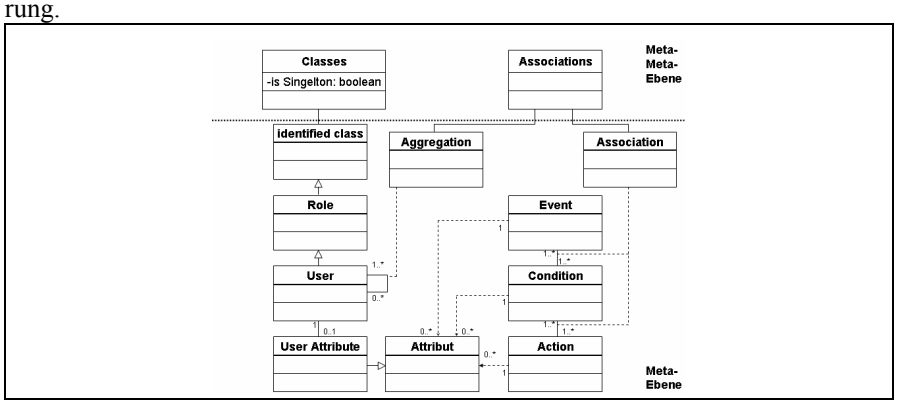

Abbildung 6: User-Profile-Meta-Modell

Dieses Metamodell enthält die Kernelemente Rolle (Role) und Benutzer (User), die durch entsprechende Attribute charakterisiert werden. *Rolle* umfasst Beschreibungselemente für die abstrakte Profilierung, *Benutzer* Elemente für die persönliche Profilierung. Zusätzlich sind die Trigger-Elemente *Ereignis*, *Bedingung* und *Aktion* enthalten, die als Operationen das dynamische Verhalten auf Basis der Attribute modellierbar machen.

Das Metamodell einer Simulationsumgebung approximiert das Verhalten von Input und Output des Simulationsmodells. Zu einem Simulationsmodell können mehrere Metamodelle erstellt werden. Ein solches Modell kann demnach immer nur speziell für die Analyse einer Ausgangsvariablen verwendet werden und nicht parallel für sämtliche Ausgangsvariablen des Simulationsmodells [Berc00, S. 2ff; Forr69]. Abbildung 7 zeigt das Metamodell zu System Dynamics. Dabei handelt es sich um ein globales Metamodell, das für sämtliche Simulationssprachen zu Grunde gelegt werden kann.

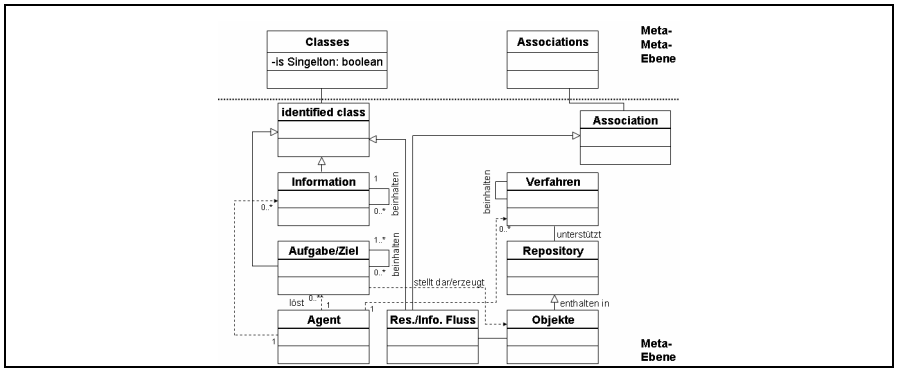

Abbildung 7: System-Dynamics-Meta-Modell (in Anlehnung an [JJPS96])

Zu sehen ist das Kernelement *Objekte*, dessen Elemente aus einem Repository abgeleitet werden. Objekte werden durch den Ressourcen- beziehungsweise Informationsfluss innerhalb des Modells beeinflusst. Es werden Verfahren angewendet, welche z. B. die Änderung eines Objektzustands repräsentieren. Dabei unterliegt das Modell einer Aufgaben-/Zielorientierung, die aus entsprechenden Informationen des Problemumfelds resultiert. Die Steuerung des Simulationsablaufes erfolgt durch einen Agenten.

In allen Metamodellen ist zu sehen, dass diese aus der MOF abzuleiten sind und somit diesem Meta-Meta-Modell genügen. Damit ist in diesem Kontext das Ziel dieser Facility erreicht, ein Integrationskonzept für unterschiedliche Metamodelle zu sein. Es lassen sich auf diesem Abstraktionsniveau Eigenschaften als Konventionen für alle Metamodelle beschreiben, so dass eine einheitliche Spezifikation innerhalb eines Informationssystems gewahrt bleibt. Die semantische Integration muss jedoch auf Ebene der Metamodelle durch die Verwendung gemeinsamer Konzepte durchgeführt werden.

4.4 Konzeption zur Integration der Meta-Modelle des Marktdateninformationssystems

Wie einleitend dargestellt, werden Inhalteräume in Anlehnung an [BeKS01, S. 20ff] konstruiert (Abbildung 8). Diese repräsentieren, nicht wie dort vorgeschlagen, einen Navigationsraum innerhalb des Informationssystems, sondern spiegeln das Informationsbedürfnis eines Anwenders wider. Das Informationsbedürfnis wird durch die Cube-Struktur, das Benutzerprofil sowie das Frühindikatorenprofil ausgedrückt.

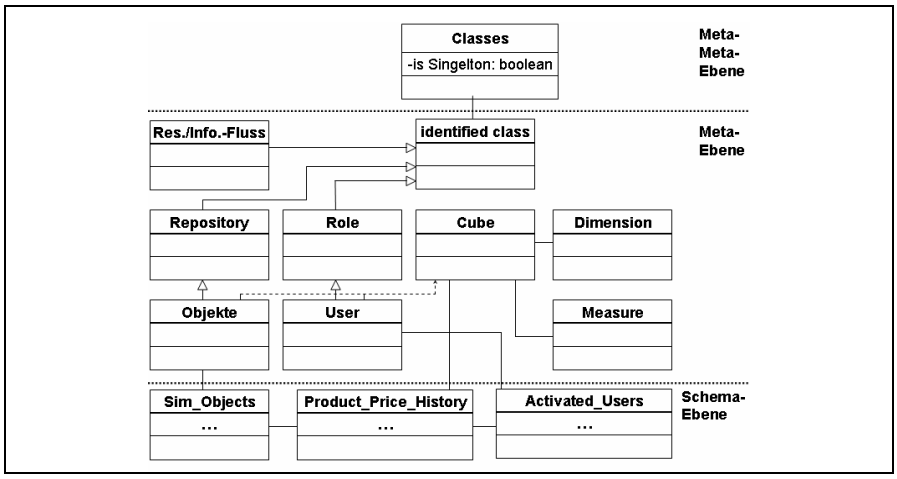

Abbildung 8: Integrierte Modellstruktur

Kernbestandteil eines ADAPT-Modells sind Cubes, die durch Dimensionen aufgespannt werden. Dies wird in der Abbildung beispielhaft durch die Struktur *Product_Price-History* gezeigt. Eine solche multidimensionale Datenstruktur repräsentiert das Informationsbedürfnis eines am System angemeldeten Anwenders (Activated_User). Eine Dimension erfasst Attribute zur Charakterisierung realer Sachverhalte. In der Struktur enthaltene Fakten bilden Urteile zu den jeweiligen Dimensionskombinationen. Qualitative Daten bilden gemäß dem MAIS-Ansatz eine weitere Dimension im Cube. Solche haben jedoch keinen direkten Bezug zu den Fakten. Für das Metamodell der Frühindikatorenprofile gilt, dass die Objekte des Repository ebenfalls reale Sachverhalte beschreiben und ebenso ein Untersuchungsobjekt für Anwender bilden. Ein Modell des Frühindikatorenprofils bezieht sich nur auf einen abgegrenzten Bereich der Realität (Sim_Objects). Diese Realität hat eine Relation zur multidimensionalen Datenstruktur, da diese auf Grund der gemeinsamen Speicherung von quantitativen und qualitativen Daten determiniert wird. Das Informationsobjekt *User* aus dem Metamodell User Profile beschreibt die Informationen, die ein Anwender automatisiert erhalten möchte. Dort sind dessen persönliche und abstrakte Charakteristika erfasst. Auch diese lassen sich einer multidimensionalen Datenstruktur zuordnen, da dort die Speicherung der Inhalte erfolgt, die ein Anwender gegebenenfalls mittels Push-Mechanismus bereitgestellt haben möchte. Entsprechend erfolgt die Integration über eine Assoziation zwischen den zentralen Modellelementen. Dies strukturiert die an den multidimensionalen Datenstrukturen orientierten Inhalteräume. Sie enthalten nicht-disjunkte Elemente. Dimensionen sind in der Regel Bestandteil mehrerer Datenstrukturen. Benutzerprofile beziehen sich auf unterschiedliche Berichtsstrukturen. Die Instanzen eines Inhalteraums stehen unabhängig von ihrer Herkunft zueinander in Beziehung. Die Beziehungen bilden die Struktur, die das Zusammenwirken der Teilsysteme realisiert.

5 Zusammenfassung und Ausblick

Im vorliegenden Beitrag wurde zunächst der Aufbau einer Architektur für ein Marktdateninformationssystem skizziert. Auch wenn eine vollständige Automatisierung zurzeit nur schwer zu erzielen ist, integriert dieser Ansatz qualitative und quantitative Daten in einer multidimensionalen Datenstruktur. Bedingt durch eine zunehmende Informationsüberflutung der Entscheidungsträger, wird eine Personalisierung der Informationsversorgung realisiert. Diese individualisierte Informationsversorgung arbeitet nicht nur vergangenheitsorientiert, sondern durch die Erweiterung um Frühindikatorenprofile zukunftsorientiert. Frühindikatoren werden in Textdokumenten identifiziert und mittels Feedback-Diagrammen potenzielle Auswirkungen aufgezeigt. Das Ergebnis erhält der Anwender entsprechend des Benutzerprofils zur Verfügung gestellt. Alle Architekturbausteine führen zu Sub-

systemen. Die einzelnen Modellierungssprachen lassen sich jedoch aus der Meta-Meta-Ebene ableiten und durch Bildung von Inhalteräumen integrieren.

Dieses Ergebnis birgt inhärente Kritik. Komplexe Modelle sind in Unternehmen in der Regel nicht wirtschaftlich zu pflegen. Dies würde umfangreiche Maßnahmen der organisatorischen Einbettung erfordern, um eine regelmäßige und konsistente Wartung und Pflege der Modelle sicherzustellen. Zusätzlich gehen Entwurf und auch Metamodellierung von Modellierungssprachen mit einer gewissen Beliebigkeit einher. Subjektive Präferenzen der jeweiligen Modellierer spielen eine bedeutende Rolle in der Modellgestaltung. Dies bringt den Nachteil mit sich, dass in verschiedenen Projekten entworfene Artefakte nur schwer miteinander in Verbindung zu bringen sind. Trotzdem ist ein multiperspektivischer Ansatz zu verfolgen. Analytische Informationssysteme werden in ihrem Leistungsangebot immer umfangreicher und komplexer. Um diese konsistent zu gestalten, bietet die Integration der Metamodelle einen Ansatz, Komplexität zu reduzieren. Damit ist es auch ein erster Schritt zu einem Unternehmensmodell, in dem sämtliche Modellierungsansätze integriert werden.

Literatur

[AbKW02] Abramowicz, W.; Kalczyński, P.; Węcel, K.: Filtering the Web to Feed Data Warehouses. Springer, London 2002.

[ArCT03] Arjona, J.; Corchuelo, R.; Toro, M.: A Knowledge Extraction Process Specification for Today's Non-Semantic Web. In: Liu, J.; Liu, C.; Klusch, M.; Zhong, N.; Cercone, N. (Hrsg.): 2003 IEEE/WIC International Conference on Web Intelligence (Proceedings, Halifax). IEEE-Publishing, Los Alamitos et al. 2003, S. 61 - 67.

[Anso80] Ansoff, H. I.: Strategic Issue Management. In: Strategic Management Journal 1 (1980), S. 131 - 148.

[BeBC+03] Becker, J.; Brelage, C.; Crisandt, J.; Dreiling, A.; Holten, R.; Ribbert, M.; Seidel, S.: Methodische und technische Integration von Daten- und Prozessmodellierungstechniken für Zwecke der Informationsbedarfsanalyse. In: Becker, J.; Grob, H. L.; Klein, S.; Kuchen, H.; Müller-Funk, U.; Vossen, G. (Hrsg.): Arbeitsberichte des Instituts für Wirtschaftsinformatik, Nr. 103, 2003.

[BeKS01] Becker, J.; Knackstedt, R.; Serries, T.: Gestaltung von Führungsinformationssystemen mittels Informationsportalen: Ansätze zur Integration von Data-Warehouse- und Content-Management-Systemen. Becker, J.; Grob, H. L.; Klein, S.; Kuchen, H.; Müller-Funk, U.; Vossen, G. (Hrsg.): Arbeitsberichte des Instituts für Wirtschaftsinformatik, Nr. 80, November 2001.

[Berc00] Berchtold, C.: Experimentieren mit Simulationsmodellen. Bericht 2000-05 der Universität der Bundeswehr, Fachbereich für Informatik, November 2000.

[Brob02] Brobst, S.: Enterprise Application Integration and Active Data Warehousing. In: von Maur, E.; Winter, R. (Hrsg.): Vom Data Warehouse zum Corporate Knowledge Center. Physica, Heidelberg, 2002, S. 15 - 22.

[Ecke03] Eckerson, W.: Smart Companies in the 21st Century: The Secrets of Creating Successful Business Intelligence Solutions. TDWI Report Series. http://www.microstrategy.com/download/files/analyst/TDWI_2003BIReport.pdf, 2003, Abruf am 2003-12-15.

[ElNa02] Elmasri, R.; Navathe, S. B.: Grundlagen von Datenbanksystemen, 3., überarbeitete Auflage. Pearson, München 2002.

[Feld02] Felden, C.: Konzept zum Aufbau eines Marktdateninformationssystems – auf der Basis interner und externer Daten. DUV, Wiesbaden 2002.

[Feld02a] Feldman, R.: Text Mining. In: Klösgen, W.; Zytkow, J. (Hrsg.): Handbook of Data Mining and Knowledge Discovery. Oxford et al. 2002, S. 749 - 757.

[Fens98] Fensel, D.: Ontologies. A Silver Bullet for Knowledge Management and Electronic Commerce. Springer, Berlin et al. 1998.

[Ferb03] Ferber, R.: Information Retrieval – Suchmodelle und Data-Mining-Verfahren für Textsammlungen und das Web. Dpunkt-Verlag, Heidelberg 2003.

[FoDu92] Foltz, P.; Dumais, S.: Personalized Information Delivery: An Analysis of Information Filtering Methods. In: Communications of the ACM, 35 (1992) 12, S. 51 - 60.

[FoKL98] Foltz, P.; Kintsch, W.; Landauer, T.: The Measurement of Textual Coherence with Latent Semantic Analysis. Discourse Process. Http:// www.knowledge-technologies.com/papers/dp2.foltz.pdf, 1998, Abruf am 2003-12-15.

[Forr69] Forrester, J. W.: Urban Dynamics. Pegasus, Cambridge 1969.

[Fran99] Frank, U.: Eine Architektur zur Spezifikation von Sprachen und Werkzeugen für die Unternehmensmodellierung. http://www.seda.sowi.uni-bamberg.de/fg510/rundbrief/Artikel%5CFran99.pdf, Abruf 2004-06-12.

[Frit02] Fritschi, H.: A Component Framework to Construct Active Database Management Systems. Dissertation http://www.ifi.unizh.ch/ifiadmin/staff/rofrei/Dissertationen/Jahr_2002/thesis_fritschi.pdf, 2002, Abruf am 2004-01-13.

[GeHK01] Gerstl, P.; Hertweck, M.; Kuhn, B.: Text Mining: Grundlagen, Verfahren und Anwendungen. In: HMD - Praxis der Wirtschaftsinformatik 38 (2001) 222, S. 38 - 48.

[Hack98] Hackathorn, R. D.: Web Farming for the Data Warehouse. Morgan Kaufmann, San Francisco 1998.

[JJPS96] Jarke, M.; Jeusfeld, M A.; Peters, P.; Szczurko, P.: Coordinating information systems engineering. Proceedings 9th Intl. Symposium on Methodologies for Intelligent Systems (ISMIS'96), Zakopane, Poland, Juni 10-13, 1996.

[Kamp02] Kamphusmann, T.: Text-Mining. Eine praktische Marktübersicht. Symposium, Düsseldorf 2002.

[KaSh00] Kashyap, V.; Sheth, A.: Information Brokering Across Heterogeneous Digital Data. A Metadata-based Approach. The Kluwer International Series on Advances in Database Systems, Nr. 20. Boston et al. 2000.

[LeRi99] Lewis, C.; Rieman, J.: Task-centered User-Interface Design – A practical Introduction. http://bmrc.berkeley.edu/courseware/cs160/fall99/Book/, 1999, Abruf am 2004-05-06.

[LiHs02] Liu, B.; Hsu, W.: User Preferences. In: Klösgen, W.; Zytkow, J. (Hrsg.): Handbook of Data Mining and Knowledge Discovery. Oxford et al. 2002, S. 467 - 474.

[Mead72] Meadows, D. H.: The limits of growth. Universe Books, New York 1972.

[Meie03] Meier, M.: Situations- und benutzerorientierte Filterung von Führungssituationen. In: Uhr, Wolfgang; Esswein, Werner, Schoop, Eric (Hrsg.): Wirtschaftsinformatik 2003, Band II, Medien - Mobile Märkte – Mobilität. Physica, Heidelberg 2003, S. 475 – 494.

[Mert97] Mertens, P.: Recommender-Systeme. In: Wirtschaftsinformatik, 39 (1997) 4, S. 401 - 404.

[OMG02] o. V.: Meta Object Facility (MOF) – Spezification, Version 1.4, November 2002. http://www.omg.org/docs/formal/02-04-03.pdf, 2002, Abruf am 2004-06-14.

[OMG01] o. V.: Common Warehouse Metamodel (CWM) – Spezification, Version 1.0, Februar 2001. http://www.omg.org/docs/ad/01-02-01.pdf, 2001, Abruf am 2004-06-14.

[PrPe01] Priebe, T.; Pernul, G.: Metadaten-gestützter Data-Warehouse-Entwurf mit A-DAPTed UML. Proceedings 5. Internationale Tagung Wirtschaftsinformatik (WI 2001) Augsburg, 2001, S. 73 - 86.

[Pull01] Pullwitt, D.: Integrating Contextual Information to Enhance SOM-based Text Document Clustering. Graduiertenkolleg Wissenspräsentation, Institut für Informatik, Universität Leipzig. Http://www.informatik.uni-leipzig.de/~pullwitt/papers/nnsi.pdf, 2001, Abruf am 2003-12-13.

[Runt00] Runte, M.: Personalisierung im Internet. DUV, Wiesbaden 2000.

[Schl00] Schlesinger, M.: ALFRED: Konzepte und Prototypen einer aktiven Schicht zur Automatisierung von Geschäftsregeln. Dissertation. Http://www.staempfli.com/digital-publications/schlesin/diss_schlesinger_ pdf_version.zip, 2000, Abruf am 2003-12-12.

[Ster00] Sterman, J. D.: Business Dynamics – Systems Thinking and Modelling for a Complex World. McGraw-Hill, New York 2000.

[Sull01] Sullivan, D.: Document Warehousing and Text Mining. Techniques for Improving Business Operations, Marketing, and Sales. Wiley, New York et al. 2001.

[VoGu01] Voß, S.; Gutenschwager, K.: Informationsmanagement. Springer, Berlin et al. 2001.