

Konzeptuelle Modellierung für modellgetriebene Decision Support Systeme

Christian Schultewolter

Universitaet Osnabrueck – Institut fuer Informationsmanagement und
Unternehmensfuehrung (IMU)

FB 9 - BWL/Management Support und Wirtschaftsinformatik

1 Motivation

1.1 Problemstellung

Entscheidungsunterstützungs- oder Decision Support Systeme (DSS) sind nach Alter (1996) ursprünglich computer-basierte Systeme, die den Benutzer bei der Identifikation und Auswahl geeigneter Entscheidungsalternativen in Situationen, zu denen keine Erfahrungen oder praktischen Lösungsansätze vorliegen, unterstützen. Die Auffassung der *Special Interest Group on Decision Support, Knowledge and Data Management Systems* der *Association for Information Systems* aus dem angloamerikanischen Raum DSS in die Klassen kommunikations-, daten-, dokumenten-, wissens- und modellgetriebene Systeme zu unterscheiden, findet heute im wissenschaftlichen Raum weltweit Unterstützung (Turban, Aronson, Liang & Sharda, 2007) und soll auch diesem Forschungsvorhaben zugrunde liegen. Modellgetriebene Decision Support Systeme (MDSS) werden oft auch synonym Spreadsheet-basierte DSS genannt, da Spreadsheet Pakete, wie z.B. Microsoft Excel, oft für die Implementierung solcher Systeme herangezogen werden (Turban et al., 2007; Holsapple und Whinston, 1996; Power und Sharda, 2007).

Ende der 1990er Jahre und zu Anfang des neuen Jahrtausends hatten Daten-getriebene Systeme den ursprünglich essentiellen Teil der Modelle größtenteils aus dem Fokus der Betrachtungen verdrängt. In jüngerer Vergangenheit lässt sich jedoch beobachten, dass die Komponente *Modell* wieder verstärkt Einfluss auf Entscheidungsprozesse, die durch DSS unterstützt werden, nimmt. Vor allem die Design- und Choice-Phasen nach Simon (1977) werden von solchen Systemen adressiert.

Panko und Sprague (1998) haben eine empirische Studie über Fehler in Spreadsheet-Modellen durchgeführt, die auch für modellgetriebene DSS relevant ist und später durch

Powell et al. überarbeitet wurde (Powell, Baker & Lawson, 2008). Die Ergebnisse zeigen, dass das Auftreten von Fehlern häufig auf eine logisch-falsche Modellierung zurückgeführt werden kann. Diese hat ihren Ursprung nicht ausschließlich in der Unkenntnis der speziellen Problemsituation und ihrer Entstehung, sondern wird durch eine einschränkende, oft wenig intuitive Modellierungskomponente forciert. Ist beispielsweise die Repräsentation expliziter Multidimensionalität nicht möglich, müssen in den meisten Fällen redundante Formeln für jegliche Ausprägungen spezifiziert werden, auf die der logisch repräsentierte Zusammenhang Anwendung finden soll. Dies mündet in einem erhöhten Risiko von Fehlern bei der Wartung, Pflege und Anwendung des Modells. Überdies hinaus können auch fehlendes Werkzeug-Wissen und unnatürliche Modellierungskonzepte die Fehlerrate negativ beeinflussen.

Analog zur Datenmodellierung könnte eine konzeptuelle Modellierungsebene Unabhängigkeit von Werkzeug-spezifischen Konzepten und Einschränkungen bieten sowie eine möglichst natürliche Modellierung ermöglichen. Ein positiver Nebeneffekt zeigt sich hier in der Berücksichtigung der klassischen DSS-Forderung die Distanz zwischen Entwickler und Entscheidungsträger zu minimieren (Sprague und Watson, 1996; Turban et al., 2007).

Wand und Weber (2002) beschreiben in ihrer Arbeit ein Framework für die weitere Erforschung von konzeptuellen Modellierungsansätzen. Es besteht aus einer Grammatik, einer Methode, einem Skript und dem Kontext, jeweils bezogen auf konzeptuelle Modellierung. Die Grammatik liefert Konstrukte und Regeln, welche die Vorschriften bei der Kombination von Konstrukten angeben, um reale Zusammenhänge zu modellieren (Wand und Weber, 2002). Sie wird in diesem Kontext durch die Klassifizierungskriterien und deren Kombination zu einer generischen Spezifikation für MDSS repräsentiert. Die Methode beschreibt wie reale Zusammenhänge durch Instanzen der Grammatik modelliert werden können (Modellierungssprache), das Skript stellt ein finales Dokument dar, welches die modellierten Zusammenhänge in Form der Sprache, erzeugt durch die Grammatik, repräsentiert. Der Kontext beschreibt die Begebenheiten und Situationen, in denen konzeptuelle Modellierung herangezogen wird.

Die Entwicklung einer konzeptuellen Modellierungsebene für MDSS wird des Weiteren durch eine nicht-prozedurale Modellspezifikation (Mayer, 1998), die als Charakteristikum eines DSS genannt wird, unterstützt. Ein konzeptueller Ansatz kann bei der Transformation nicht-prozedural modellierter Zusammenhänge in korrekte, prozedurale Rechenschritte helfen, indem er konsistente Kriterien für die Klassifizierung verschiedener Beziehungstypen zur Verfügung stellt. Er trägt dadurch zum Entstehen eines Frameworks bei, das während des Transformationsprozesses als Checkliste herangezogen werden kann.

Um einen Eindruck des Anwendungsbereiches zu erhalten, wird im weiteren Verlauf des Artikels eine Übersicht über typische Eigenschaften modellgetriebener DSS gegeben, für die die konzeptuelle Ebene entworfen werden soll. Nachdem die Zielsetzung konkretisiert worden ist, werden drei Klassifizierungskriterien vorgeschlagen, die eine systematische Repräsentation von Beziehungen innerhalb eines Modells ermöglichen. Im Anschluss untersucht Kapitel vier bekannte Modellierungsansätze aus verwandten Themengebieten, um eine Eignung dieser für die Umsetzung einer konzeptuellen Modellierungssprache darzulegen. Schließlich wird ein kurzer Ausblick über mögliche Szenarien und Forschungsoptionen für das weitere Vorhaben umrissen.

1.2 Modellgetriebene Decision Support Systeme

Die Hauptaufgabe modellgetriebener DSS besteht in der dynamischen, mehrfachen Analyse oft vektorieller Formeln oder Beziehungen, die zwischen mehreren Teilen oder Variablen in einem diskreten, multidimensionalen Modell bestehen (Power und Sharda, 2007). Im Gegensatz hierzu richten datengetriebene Systeme ihr Hauptaugenmerk auf die Konsolidierung und zeitspezifische Organisation großer Datenmengen, z.B. in Data Warehouses (March und Hevner, 2007). Das Leistungsvermögen und das Potenzial modellgetriebener DSS werden von der Mächtigkeit des Funktionsspektrums definiert, welches dem Anwender bei der Modellierung zur Verfügung steht. Demnach also den Algorithmen und mathematischen Formeln (Power und Sharda, 2007; Sprague und Watson, 1996), die für die Darstellung realer Situationen zur Verfügung stehen. Es sei jedoch erwähnt, dass MDSS nicht der einzige Typ von DSS sind, welche diese Form von Modellen nutzen. Datengetriebene Systeme beinhalten ebenfalls mathematische Modelle, z.B. um abgeleitete Kennzahlen in einem On-Line Analytical Processing (OLAP-)System zu berechnen (Koutsoukis, Mitra & Lucas, 1999). Der Unterschied zwischen den beiden Systemen wird deutlich, wenn man die Zugriffsmöglichkeiten näher betrachtet, die dem Anwender während der Nutzung des Systems auf die Formeln und Beziehungen innerhalb des Modells zur Verfügung stehen. In datengetriebenen Systemen kann deren Manipulation lediglich im Vorfeld der Erstellung einzelner Modell-Instanzen (z.B. OLAP-Würfel), beispielsweise durch Änderung der ETL-Prozesse, stattfinden. So ist es zwar möglich mehrere Szenarien innerhalb eines OLAP-Würfels abzulegen, jedoch ist es im Nachhinein nicht möglich beispielsweise die vordefinierten Dimensionen um Weitere zu ergänzen (March und Hevner, 2007) und somit zwei beliebig unterschiedliche Modell-Instanzen miteinander zu vergleichen. Modellgetriebene DSS stellen dem Anwender hingegen die gesamte Modelllogik zur Verfügung und ermöglichen dadurch unterschiedliche Analysen mit beliebigen Modellparametern, ohne die Einschränkung vordefinierter Fixwerte. Diese Unterschiede führen zu einem typischen Funktionsspektrum je DSS-Typ. Für MDSS sind hier primär Zielwertsuchen (Goal-seeking), What-if und Sensitivitätsanalysen zu erwähnen.

Bei der Zielwertsuche wird ein bestimmter, von anderen Modellkomponenten abhängiger Wert (bezogen auf Spreadsheet-basierte Systeme: ein Zellwert), derart abgeändert, dass ein gewünschtes Ergebnis erzielt wird (Coles und Rowley, 1996). Systemseitig wird nun ermittelt welche Zellen in einer direkten oder indirekten Beziehung zu dem geänderten Wert stehen. Durch Rückwärtsrechnungen, Formelumstellungen und bei diversen Werkzeugen Anwenden gezielter Verteilungsstrategien – z.B. ist denkbar, dass einige Werte unverändert bleiben sollen oder einzelne Formelbestandteile überproportional gewichtet werden (Coles und Rowley, 1996) – werden deren Werte unter den geänderten Voraussetzungen neu berechnet. In einer What-if Analyse werden hingegen in der Regel Werte von unabhängigen Zellen (Rohdaten) oder aber Formelverknüpfungen variiert, um im Weiteren die Auswirkungen der Änderungen auf das Modell und vor allem abhängige Zielwerte zu beobachten (Holsapple, 2008). Der Hauptunterschied einer Sensitivitätsanalyse findet sich in den ermittelten Kennzahlen. Diese beschreiben keine resultierenden Werte oder Auswirkungen leichter Daten- oder Strukturvariationen für ein Modell, sondern ermitteln für zuvor bestimmte Parameter diejenigen Bereiche, in denen eine Wertänderung keine Auswirkung auf abhängige Teile des Modells hat (Werners, 2008; Insua, 1990). Die Bestimmung erfolgt in der Regel mittels komplexer Algorithmen (z.B. lineare Programmierung), mit deren Hilfe diese Einfluss-Elastizitäten bestimmt werden können (Turban et al., 2007).

1.3 Zielsetzung

Das Gestaltungsziel der Arbeit ist das Entwickeln einer konzeptuellen Modellierungsebene einschließlich einer Modellierungssprache, die den Anwender beim Abbilden logischer Zusammenhänge in einem modellgetriebenen DSS unterstützen kann. Es soll ein Vorschlag erarbeitet werden, der das Einfügen einer solchen Ebene in die Architektur modellgetriebener DSS thematisiert.

Um dieses Ziel zu erreichen, muss sichergestellt werden, dass die Modellierungssprache sämtliche zu modellierende Fälle abbilden kann. Des Weiteren sollte eine möglichst standardisierte technische Implementierung die Interaktion mit bestehenden DSS-Werkzeugen unterstützen. Hierfür sind vorhandene Ansätze multidimensionaler Modellierung auf ihre Eignung zu überprüfen.

Aufbauend auf diesen Gegebenheiten wird im nächsten Kapitel ein Vorschlag von Kriterien zur Klassifizierung unterschiedlicher Beziehungen innerhalb dieser Systeme ausgearbeitet. Für ein leichteres Verständnis wird jede Beziehung durch ein vereinfachtes Beispiel veranschaulicht und durch eine kleine Skizze visuell unterstützt.

2 Klassifizierungskriterien von Modellbeziehungen

2.1 Basisannahmen

Zentrales Element der vorgeschlagenen generischen Spezifikation ist die (formelmäßige) Beziehung zwischen Modellzellen innerhalb eines diskreten, multidimensionalen Raums. Jede Dimension eines Raums soll generell die gleichen Typen arithmetischer Verknüpfungen anbieten können. Demzufolge sollen Dimensionen, die eine spezielle Aufgabe wahrnehmen oder spezielle Funktionen anbieten, vermieden werden. Dadurch ist es möglich auch große Modelle schnell und leicht verstehen und mit angemessenem Aufwand pflegen zu können. Letztlich kommt diese Forderung der flexiblen Anwendungen von DSS-Funktionen zu gute. So muss beispielsweise sichergestellt werden, dass für eine konsolidierende Ausprägung Jahressumme Werte anderer Dimensionen summiert (z.B. Mengen oder Umsätze) oder als Durchschnitt (z.B. Quoten oder Preise) berechnet werden. Eine ähnliche Forderung findet sich bereits in den zwölf Regeln für OLAP-Systeme von Edgar F. Codd (Regel 6 - Generische Dimensionalität) (E. Codd, S. Codd & Salley, 1993). Diesem Ansatz folgend ist der Ausdruck Dimension im Weiteren als generischer Begriff zu verstehen. Das bedeutet, dass Kennzahlen und Dimensionen, im Gegensatz zu anderen mehrdimensionalen Modellierungsansätzen verwandter Themengebiete, keine streng getrennten Bestandteile eines Modells mit unterschiedlichen Bedeutungen darstellen (wie z.B. im SDWM, vgl. Kapitel 3.1), sondern Kennzahlen auch qualifizierende Dimension repräsentieren können. Des Weiteren ist zu erwähnen, dass eine Verknüpfung, die eine bestimmte Logik ausdrückt, lediglich einmal definiert und der Modellierungsprozess nicht-prozedural sein sollte, um eine konsistente und nicht redundante Modellierung gewährleisten zu können.

2.2 Kriterien zur Klassifizierung

Die Konstruktion einer generischen Spezifikation möglicher Modellbeziehungen bildet die Basis für eine konzeptuelle Modellierungsebene. Dem Framework für die weitere Erforschung von konzeptuellen Modellierungsansätzen von Wand und Weber (2002) folgend, müssen die Kriterien für eine generische Spezifikation von Modellbeziehungen in MDSS die Modellierung sämtlicher realer Situationen und Zusammenhänge ermöglichen. Die Arbeit von Green (1997) erarbeitet für dieses Framework die beiden Kriterien minimale ontologische Überschneidung (MOO - minimal ontological overlap) und maximale ontologische Vollständigkeit (MOC - maximum ontological completeness), welche durch adäquaten Einsatz der Grammatik erreicht werden sollen. MOO drückt aus, dass ein Konstrukt nicht durch andere grammatikalische Konstrukte oder deren Kombination ausgedrückt werden kann. Hierdurch verringert sich die Wahrscheinlichkeit, dass konfliktäre ontologische Repräsentationen eines realen Zusammenhangs auftreten kön-

nen. Bezogen auf dieses Paper bedeutet dies also, dass es für die Modellierung realer Situationen lediglich eine korrekte Möglichkeit (Kriterien-Kombination) gibt. Maximale ontologische Vollständigkeit wird erreicht, wenn durch die Kombination der vorhandenen grammatikalischen Konstrukte sämtliche reale Phänomene mittels der Grammatik repräsentiert werden können. Gleichzeitig erhöht sich die Wahrscheinlichkeit, dass eine Domäne komplett dargestellt ist.

Jede mögliche Situation, die modelliert werden soll, wird durch ein betriebswirtschaftliches Beispiel veranschaulicht. Diese fokussieren vor allem die Darstellung der dimensionalen Differenzierung von Modellbeziehungen sowie die Herausforderung der Transformation von nicht-prozedural modellierten Zusammenhängen in automatisiert ausführbare Berechnungsschritte (Rieger, 1993). Aufgrund dieser Zusammenhänge und der Möglichkeit sämtliche reale Zusammenhänge modellieren zu können, müssen die Klassifizierungskriterien in der Lage sein die charakteristischen, multidimensionalen Strukturen und sich daraus ergebenden Konsequenzen für die Transformation zu unterscheiden. Multidimensionale Räume in Kombination mit vektoriiellen Formeln führen innerhalb Spreadsheets-basierter Systeme in der Regel zu Überschneidungen von Formeln für gewisse Zellbereiche, die bestimmte Formelreihenfolgen bedingen, und speziellen Berechnungsreihenfolgen beteiligter Zellen. Diese beiden Reihenfolgen und deren Unterscheidung sind für das weitere Verständnis eminent wichtig, da sie aus unterschiedlichen Zusammenhängen resultieren. Sie haben ihren Ursprung in speziellen Eigenschaften, die von den Klassifizierungskriterien auf angemessene Art und Weise dargestellt werden müssen. Werden diese Forderungen berücksichtigt, erfüllt eine derartige Spezifikation die Kriterien nach Green sowie Wand und Weber zwangsläufig ebenso. Folgende binäre (zwei mögliche Ausprägungen) Kriterien, welche in ihrer Grundstruktur auf Rieger (1993) zurückgehen, werden vorgeschlagen: *Grad*, *Homogenität* und *Simultaneität* von Dimensionen, die an einer Beziehung beteiligt sind (vgl. Abbildung 1).

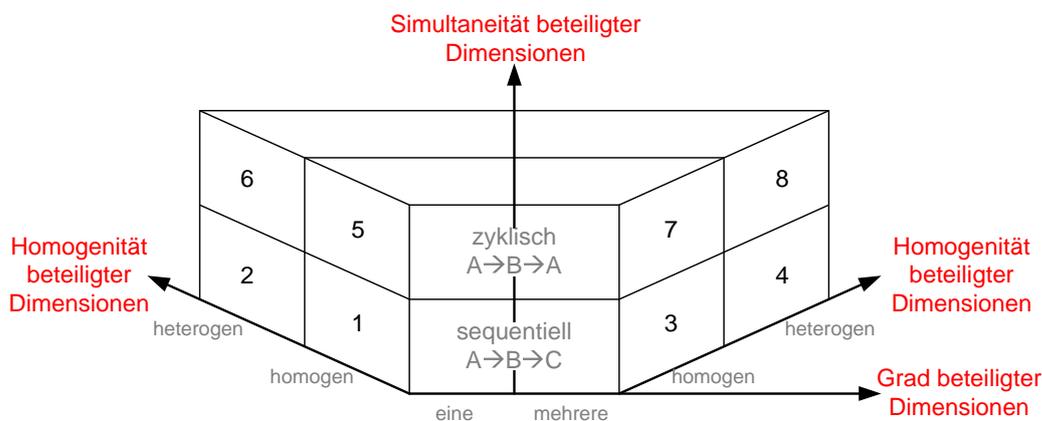


Abbildung 1: Fallklassifizierung modellgetriebener DSS

Folgendes Beispiel erläutert kurz die Wirkung der Ausprägung *Grad*: In einem Modell sind die beiden Dimensionen Kennzahlen und Zeit enthalten. Auf einer Kennzahlen Dimension befinden sich die beiden Formeln „Umsatz = Menge * Preis“ und „Gewinn = Umsatz – Kosten“. Auf der Dimension Zeit ist lediglich eine Summenformel, die die zwei Ausprägungen „2009“ und „2010“ in einer Ausprägung „Summe“ addiert, vertreten. Bei intra-dimensionalen Formeln ist die Reihenfolge der Dimensionen hier irrelevant, stattdessen spielt nur die Sortierung der Zellen innerhalb der Dimension eine Rolle. Für inter-dimensionale Referenzen muss die Reihenfolge der Dimensionen für alle betroffenen Zellen berücksichtigt werden. Dies ist der Fall, da die Werte referenzierter Zellen, die mit Formeln einer anderen Dimension belegt sind, als der an einer bestimmten Stelle des Berechnungsprozesses aktuellen, bereits vorliegen müssen. Diejenigen Zellen, die nicht aus der Kreuzung mit den Dimensionsausprägungen „2009“ oder „2010“ der Zeit Dimension entstehen, sind zugleich mit der Summenformel aus der Zeit Dimension belegt. Für diese muss ebenfalls die Reihenfolge der Dimensionen, d.h. der Berechnungsformeln aus unterschiedlichen Dimensionen, berücksichtigt werden, um ein korrektes Modell zu erhalten. Generell werden konsolidierende Formeln, wie im vorliegenden Beispiel die Summenformel der Zeit Dimension, eher am Ende eines Berechnungsprozesses platziert, es gibt jedoch Ausnahmen.

Das Kriterium *Homogenität* gibt an, ob die logische Struktur eines Modellierungsproblems die Bildung von Intervallen erfordert. Mögliche Ausprägungen sind homogen, falls keine Fallunterscheidungen nötig sind und heterogen für Fälle, in denen mehrere Formeln unterschieden werden müssen (Intervalle, vgl. Abbildung 1). Liegen verschiedene Intervalle für eine Dimensionsausprägung vor, so müssen intra- und inter-dimensionale Strukturen unterschieden werden. Bei intra-dimensionalen Formeln kann lediglich ein Wert innerhalb derselben Dimension für die Unterscheidung der Formeln verantwortlich zeichnen. Im inter-dimensionalen Fall ist es möglich, dass Ausprägungen weiterer Dimensionen die Anwendung intervallspezifischer Formeln bedingen. Bei mehreren, unterschiedlichen Formeln für eine Dimensionsausprägung kann man sich eine Intervallbildung auch als Partitionierung des Modellraumes vorstellen.

Das dritte und letzte Kriterium *Simultanität* gibt an, ob simultane Strukturen innerhalb von Formeln oder Referenzen vorliegen. Die möglichen Werte werden hierbei durch zyklisch und sequentiell dargestellt. Simultanität besteht, wenn zwei oder mehr Dimensionsausprägungen miteinander zur selben Zeit interagieren, so dass ein Kreisschluss entsteht (Zeit-diskret). In betriebswirtschaftlichen Problemen spielt Simultanität oft eine wichtige Rolle. Je mehr Details bei der Modellierung berücksichtigt werden, desto realistischer ist das Modell der realen Zusammenhänge, da weniger Annahmen getroffen werden müssen. Die Wahrscheinlichkeit eines Auftretens simultaner Strukturen erhöht sich mit dem De-

taillierungsgrad der modellierten Zusammenhänge, da bei mehreren Größen in der Regel auch mehrere Interaktionen innerhalb eines Modells stattfinden (ansonsten würde das Berücksichtigen einer Größe wenig Sinn machen). Simultanität kann sich auch hinter einer Kette formelmäßiger Zusammenhänge verbergen. Auch in solchen Fällen muss jedoch eine gesonderte Rechnung durchgeführt werden, um ein valides Modell zu erhalten.

2.3 Szenarien

Der folgende Abschnitt erläutert acht Modellierungsfälle anhand von typischen betriebswirtschaftlichen Beispielen. Die acht Fälle ergeben sich aus der kartesischen Kombination der Ausprägungen der Klassifizierungskriterien untereinander, die im vorangegangenen Kapitel vorgestellt wurden (drei Kriterien á zwei Ausprägungen). Die Beispiele sind stark vereinfacht dargestellt und auf zwei Dimensionen limitiert, da dies eine einfache Darstellung der wesentlichen, bedeutsamen Zusammenhänge ermöglicht. Des Weiteren würde das Einführen weiterer Dimensionen in dem hier präsentierten Kontext keinen Erkenntnisgewinn mit sich bringen. Die charakteristischen Wirkungen werden für jeden Fall mittels eines Ursache-Wirkungs-Diagramms skizziert. Ein roter Pfeil gibt hierbei die Wirkungsrichtung der Formeln für jeden spezifischen Fall an. Ist dieser diagonal angeordnet, liegt eine inter-dimensionale Beziehung und wenn er vertikal verläuft, liegt eine intra-dimensionale Beziehung vor. Zwei gegenläufige Pfeile stellen wiederum simultane Strukturen dar und eine gestrichelte Linie repräsentiert heterogene Beziehungen, für die eine Fallunterscheidung hinsichtlich der Formeln vorgenommen werden muss. Zweitrangige oder analoge Wirkungen werden innerhalb der Diagramme nicht berücksichtigt.

Im Folgenden werden zunächst die Fälle 1 bis 4 (vgl. Abbildung 2) behandelt, die sequentielle Strukturen kennzeichnen. Danach werden die Fälle 5 bis 8 (vgl. Abbildung 3), die simultane Strukturen aufweisen, näher erläutert.

Fall 1:

Fall 1 kennzeichnet intra-dimensionale Beziehungen, welche ohne Fallunterscheidung und ausschließlich sequentielle Strukturen aufweisen. Das bedeutet, dass eine intra-dimensionale Formel „Umsatz = Menge * Preis“ auf jede Zelle angewandt wird, die durch Kreuzung mit der Dimensionsausprägung „Umsatz“ entsteht.

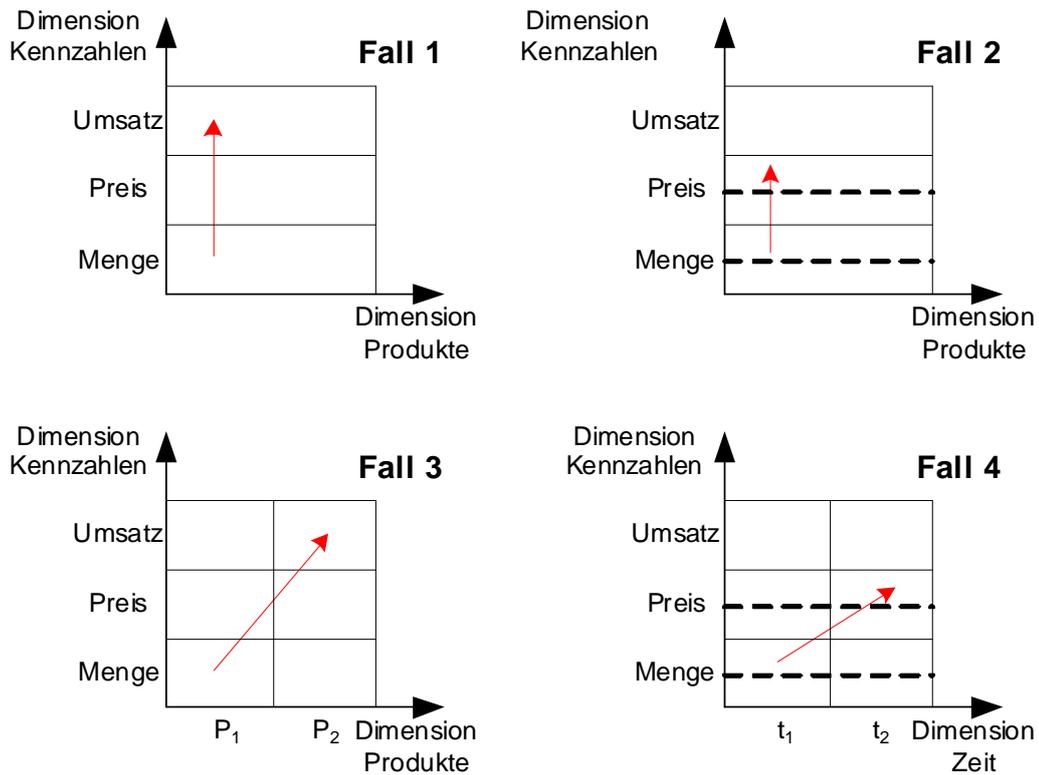


Abbildung 2: Falle 1 bis 4 mit sequentieller Struktur

Die korrekte Reihenfolge der Berechnungsschritte kann durch einfaches Umsortieren der Zellen bzw. ein Sortieren der Formeln fur Zellen, die mit mehr als einer Formel belegt sind, erreicht werden. Fur alle Zellen, die lediglich mit intra-dimensionalen Formeln belegt sind, ist die Reihenfolge der Dimensionen irrelevant (vgl. Beispiel fur die Wirkung des Klassifizierungskriteriums Grad aus Kapitel 3.2).

Fall 2:

Zusatzlich zu Fall 1 werden in diesem Fall Intervalle, das heit heterogene Formeln eingefuhrt. Spezifische Werte einer beliebigen Dimensionsauspragung sorgen bei Beziehungen dieser Art dafur, dass unterschiedlich definierte Formeln fur die Berechnung ein und derselben Dimensionsauspragung verantwortlich zeichnen. In Bezug auf das Beispiel aus Fall 1 konnte ein Rabattsatz, der ab einer bestimmten Menge gewahrt wird, diesen Fall darstellen. Die Formeln konnten wie folgt gestaltet sein: „Umsatz = Menge * Preis * (1 – Rabattsatz)“. Fur eine Menge, die kleiner als 100 ist, wurde der Rabattsatz automatisch „0“ sein und fur Mengen groer 100 durch eine Formel „Menge / 10.000“ berechnet werden.

Für die korrekte Berechnungsreihenfolge muss der Umsatz bei Mengen größer „100“ immer nach dem Rabatt berechnet werden. Für Mengen kleiner „100“ fällt diese Sperre weg, da die Formel für den Umsatz an beiden Stellen des Berechnungsprozesses das gleiche Ergebnis zur Folge hätte.

Fall 3:

Während in Fall 1 und 2 ausschließlich intra-dimensionale Referenzen vorlagen beschreibt Fall 3 einen „echten“ multidimensionalen Raum, in dem eine Formel mindestens zwei Dimensionen miteinander in Relation setzt. Eine eingeschränkte Referenz bezüglich bestimmter Werte (Intervalle) liegt jedoch nicht vor.

Bei der Transformation nicht-prozedural modellierter Strukturen in prozedurale Berechnungsschritte muss die Reihenfolge, nach der die Dimensionen abgearbeitet werden, beachtet werden. Wird eine Dimensionsausprägung innerhalb einer Formel, die einer weiteren Dimension entstammt, referenziert, muss die betreffende Zelle zuvor berechnet worden sein. Es wird schnell klar, dass es nicht machbar ist, Zellen einer ganzen Dimension nacheinander zu berechnen, sondern eine bestimmte Reihenfolge unabhängig von Dimensionen bei der Berechnung der Zellen hergestellt werden muss.

Ein Beispiel für diesen Typ von Beziehung stellen Formeln für die Abbildung eines mehrstufigen Produktionssystems dar. Es muss sichergestellt werden, dass nicht mehr (Zwischen-)Produkte zur Herstellung eingeplant werden, als es die jeweils vorhergehende Stufe zulässt. So kann zum Beispiel in einer Tischlerei nicht mit 100 Tischen (vgl. Abbildung 2 Fall 3 Größe P_2) geplant werden, sofern lediglich 100 Tischbeine (vgl. Abbildung 2 Fall 3 Größe P_1) für den nächsten Produktionslauf vorliegen.

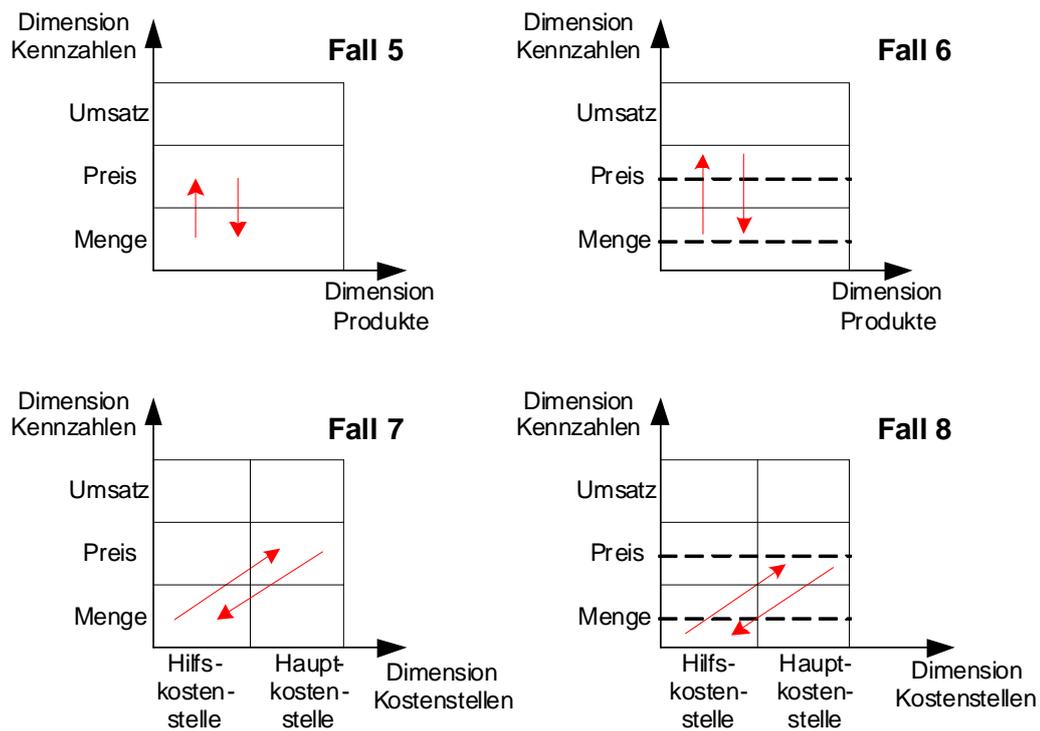


Abbildung 3: Fälle 5 bis 8 mit simultaner Struktur

Fall 4:

Fall 4 als letzte sequentielle Beziehung ermöglicht zusätzlich zu Fall 3 eine Fallunterscheidung (Intervalle) für bestimmte Werte von Dimensionsausprägungen. Im Vergleich zum Beispiel aus Fall 2 würde die Absatzmenge einer vorangegangenen Periode über die Möglichkeit der Gewährung eines Rabattes entscheiden. Dies hätte zur Folge, dass in der Formel für den Rabattsatz eine Zeitdimension zugegriffen werden müsste.

Die folgenden Fälle 5 bis 8 (vgl. Abbildung 3) sind durch simultane Berechnungsstrukturen charakterisiert. Eine simultane Formel muss in einer gesonderten Nebenrechnung, für die es mehrere Strategien gibt, aufgelöst werden. Häufig wird auf Modelle Linearer Programmierung zurückgegriffen, da diese in der Regel auf einfache Art und Weise sehr nah an die optimale Lösung herankommen.

Fall 5:

Der erste Fall mit simultanen Strukturen berücksichtigt lediglich eine Dimension und keine Fallunterscheidungen für die Formel einer Dimensionsausprägung.

Nimmt man das Beispiel aus Fall 1 bezüglich der einfachen Umsatzberechnung, werden in diesem Fall nun eine Angebots- und eine Nachfragefunktion abgebildet. Diese beeinflussen sich jeweils gegenseitig, indem die Menge vom Preis (Angebotsfunktion) und der Preis von der Menge (Nachfragefunktion) abhängen.

Beim Etablieren einer korrekten Berechnungsreihenfolge bezüglich Zellen und Formeln muss jede Zelle, die direkt von der simultanen Formel abhängt, in einer Nebenrechnung berechnet werden. In dem Beispiel wären dies die beiden Zellen „Preis“ und „Menge“ des abgebildeten Produkts. Die Nebenrechnung müsste nun an einer bestimmten Stelle des gesamten Berechnungsprozesses platziert werden. Diese hängt wiederum von der Sortierung der restlichen Zellen und dem Zusammenspiel dieser mit den simultanen Größen ab. Für Zellen, die mit mehr als einer Formel belegt sind, die aus unterschiedlichen Dimensionen stammen, ist zudem die Formelsortierung (Reihenfolge der Dimensionen) relevant.

Fall 6:

Fall 6 führt die Berücksichtigung von unterschiedlichen Formelalternativen für eine Dimensionsausprägung im simultanen Kontext ein. Zusätzlich zu Fall 5 wären die Angebots- bzw. Nachfragefunktion für ein bestimmtes Mengenintervall (Nachfragefunktion) bzw. ein bestimmtes Preisintervall (Angebotsfunktion) völlig unelastisch. Ab einer bestimmten Menge bzw. einem bestimmten Preis würden sich die Funktionen jedoch auf ähnliche Weise, wie in Fall 5 beschrieben, beeinflussen. Dies hat zur Folge, dass die unterschiedlichen Intervallformeln, sowohl für die Menge als auch für den Preis, an unterschiedlichen Stellen des Berechnungsprozesses platziert werden würden, da die simultanen Zusammenhänge wiederum in einer Nebenrechnung gelöst werden müssen.

Fall 7:

Dieser simultane Zusammenhang verbindet Ausprägungen unterschiedlicher Dimensionen miteinander, unterscheidet hierbei jedoch keine Formelalternativen (homogen).

Die innerbetriebliche Leistungsverrechnung ist ein klassisches Beispiel hierfür. Eine Hilfskostenstelle ist einer Hauptkostenstelle zugeordnet. Auf der einen Seite konsumiert die Hilfskostenstelle Leistungen der Hauptkostenstelle, auf der anderen Seite stellt sie der Hauptkostenstelle Leistung zur Verfügung. Die benötigte Leistung der Hilfskostenstelle beeinflusst den Preis der Hauptkostenstelle und umgekehrt.

Um in diesem Fall eine korrekte Berechnungsreihenfolge der Zellen und Formeln herzustellen müssen wiederum die Reihenfolge der Dimensionen beachtet und die Nebenrechnung zur Auflösung der simultanen Größen an eine spezifische Stelle des Gesamtprozesses eingegliedert werden.

Fall 8:

Der achte und letzte Fall deckt die Möglichkeit von simultanen Formeln in einem multi-dimensionalen Raum ab, die – im Gegensatz zu Fall 7 – nicht in gleicher Weise für alle Werte der betroffenen Dimensionsausprägung gelten. Erweitert man das Beispiel der innerbetrieblichen Leistungsverrechnung aus Fall 7 dahingehend, dass Kapazitätsgrenzen die Menge an Leistung beschränken, die eine Hilfskostenstelle einer Hauptkostenstelle zur Verfügung stellt, ergibt sich als Konsequenz, dass die Hauptkostenstelle selbige Leistung von außen zukaufen muss. Für einige Werteausprägungen von Elementen der Dimension Kennzahlen gilt also eine unterschiedliche Berechnungsvorschrift (einmal „intern“ und einmal „extern“).

Die unterschiedlichen Intervallformeln müssen an unterschiedlichen Stellen des Berechnungsprozesses platziert werden da die simultanen Zusammenhänge sich auf ein Intervall beziehen und wiederum in einer Nebenrechnung gelöst werden müssen. Diese muss in den gesamten Berechnungsprozess eingegliedert werden und zudem ist die Reihenfolge der Dimensionen aufgrund der inter-dimensionalen Beziehungen relevant.

Im Folgenden werden nun bekannte Ansätze konzeptueller Modellierung kurz vorgestellt und auf ihre Potenziale bei der Realisierung einer konzeptuellen Modellierungsebene in MDSS vor dem Hintergrund der gerade vorgestellten Fallklassifizierung untersucht.

3 Bekannte Ansätze konzeptueller Modellierung

3.1 Semantisches Data-Warehouse Modell (SDWM)

Das semantische Data-Warehouse Modell (Böhnlein und Ende, 2001) ist ein an der Universität Bamberg entwickelter Ansatz, dessen Intention die Bereitstellung einer rein konzeptuellen Betrachtungsweise multidimensionaler Datenstrukturen im OLAP- bzw. Data-Warehouse Umfeld ist. Grundsätzlich liegt dem Ansatz das Konzept der Bildung von Sichten zu Grunde. Die Autoren bezeichnen quantitative Daten als Kennzahlen und qualitative Aspekte als Dimensionen.

Die Dimensionssicht wird durch die Bestandteile Dimension, Dimensionshierarchiestufe, Aggregationsbeziehung und Dimensionsschnittstelle charakterisiert. In einer Dimension sind zumindest zwei Dimensionshierarchiestufen und eine Aggregationsbeziehung vorhanden. Durch die Aggregationsbeziehung werden Dimensionshierarchiestufen miteinander verbunden. Die Dimensionsausprägungen sind einzelnen Hierarchiestufen zugeordnet und bilden von der niedrigsten zur höchsten Stufe eine Verdichtung der Daten über eine Aggregation ab. Eine Dimensionsschnittstelle bildet eine Verknüpfung von Dimensionen

mit einer einzelnen Basiskennzahl ab. Dies stellt die Wiederverwendbarkeit von Dimensionen sicher.

Hauptbestandteile der Kennzahlensicht sind die Elemente Basiskennzahl, abgeleitete Kennzahl, Dimensionsschnittstelle und Dimension. Das Kombinieren von Dimensionen und Kennzahlen als zentrale Eigenschaft multidimensionaler Datenstrukturen steht hier im Fokus. Basiskennzahlen, welche über eine eindeutige Dimensionsschnittstelle mit einer Dimension verknüpft werden, können über gerichtete Kennzahlbeziehungen zu abgeleiteten Kennzahlen kombiniert werden. Eine abgeleitete Kennzahl besitzt zur Spezifikation ihrer Zusammensetzung ein Attribut Berechnungsvorschrift. Die Möglichkeit nicht-additive, semi-additive und uneingeschränkt additive Verdichtungen innerhalb von Dimensionen abzubilden, wird durch ein Attribut Aggregierbarkeit, welches der Dimensionsschnittstelle zugeordnet ist, realisiert. Ebenfalls findet sich dort ein Attribut Dimensionslinie, welches die Granularität einer Basiskennzahl angibt, d.h. bis zu welcher Hierarchiestufe einer Dimension eine Verfeinerung möglich ist.

Bezogen auf den Kontext modellgetriebener DSS ergeben sich aus diesen Strukturen Einschränkungen der Modellierungsmächtigkeit. So ist es z.B. nicht möglich, Teilmengendifferenzierte Kennzahlenspezifikationen innerhalb einer Dimension zu definieren (keine Intervalle bzw. heterogene Dimensionen in Bezug auf Referenzen möglich). Des Weiteren ergibt sich, dass andere Beziehungen als Aggregationen zwischen zwei Dimensionsausprägungen (auf derselben Hierarchiestufe) nicht mittels SDWM abzubilden sind. Simultane Referenzen, welche für die Mächtigkeit modellgetriebener DSS oft relevant sind, werden in diesem System nicht gesondert berücksichtigt. Eine prinzipiell denkbare SDWM-Erweiterung soll hier nicht weiter verfolgt werden, da dies angesichts des Sichtenkonzepts zu redundanten Modellierungskonzepten führt.

3.2 Multidimensional Modeling Language (MML)

Die Multidimensional Modeling Language (Herden und Harren, 2004) ist eine an der Universität Oldenburg entwickelte Sprache, zur konzeptionellen Modellierung multidimensionaler Data-Warehouse-Schemata. Grundsatz ist die Differenzierung von Daten, dem zugehörigen multidimensionalen Kontext und Elementen für die Beschreibung von Struktureigenschaften. Die MML wurde mittels einer multidimensionalen Erweiterung der Unified Modeling Language (UML) realisiert und nutzt typische Konstrukte aus der objektorientierten Programmierung, wie z.B. Vererbung oder Generalisierung.

MML besteht aus verschiedenen Konstrukten, die in die fünf Bereiche Daten-Elemente, Hilfsmetaklassen, Multidimensionaler Kontext, Allgemeine Verbindungen und Properties unterteilt sind. Ursprung eines jeden MML-Konstrukts ist die abstrakte Metaklasse MMLElement, in der jeder Klasse eine eindeutige Bezeichnung zugeordnet wird. Der

Bereich Daten-Elemente entspricht einem Datentyp und beschreibt, ob es sich bei einem Daten aufnehmenden Element um einen elementaren oder komplexen Typ handelt. Im Bereich Hilfsmetaklassen werden die Eigenschaften der Vererbung einer Klasse gesteuert und der Bereich Multidimensionaler Kontext führt die bekannte Unterscheidung in Fakten und Dimensionsdaten ein. Mögliche Verknüpfungen zwischen Schemaelementen werden in den Bereichen Allgemeine Verbindungen und Properties geregelt. Während erstgenannter typische objektorientierte Konstrukte, wie z.B. Assoziation und Komposition zur Verfügung stellt, dient der Bereich Properties der Darstellung multidimensionaler Sachverhalte, wie z.B. dem Aufbau von Hierarchien.

Bezogen auf ein modellgetriebenes DSS folgen hieraus eingeschränkte Möglichkeiten bei der Modellierung bestimmter Sachverhalte. So ist es zwar für eine bestimmte Grundgesamtheit von Elementen möglich mittels eines „SharedRollUp“ aus dem Bereich Properties intervallspezifische Berechnungsvorschriften zu definieren, eine Betrachtung von Dimension als generischer Begriff, der auch Faktenelemente als qualifizierende Dimension zulässt, ist jedoch nicht möglich. Weitere Einschränkungen liegen bei der Berücksichtigung von Besonderheiten simultaner Referenzen vor. Diese bleiben gänzlich unberücksichtigt, was eine gesonderte Behandlung im Weiteren erschwert.

4 Weiteres Vorgehen

Aus wissenschaftstheoretischer Sicht sollen die bislang vorgestellten Ergebnisse im Sinne des konstruktionsorientierten Ansatzes in Verbindung mit deduktiven Elementen zur Erlangung neuer Erkenntnisse fortgeführt werden. Dieser Ansatz bietet sich hier vor allem deshalb an, da eine Situation gegeben ist, zu der bisher keine oder nur weniger zufriedenstellende Lösungsansätze vorhanden sind, sich gleichwohl aber eindeutige Anforderungen spezifizieren lassen (Frank, 2007). Diese Anforderungen konnten durch die, auf Basis von aus der Literatur gewonnenen Kriterien, hergeleitete Fallklassifizierung für modellgetriebene DSS definiert werden. Auf Basis dieser Vorarbeiten werden folgende Forschungsaspekte für konkrete Schritte im Rahmen des Dissertationsprojekts gesehen:

Die anhand von singulären Beispielen veranschaulichten Modellierungsfälle sollen in einem nächsten Schritt zu einem Referenzmodell weiterentwickelt werden. Im Folgenden soll diese Referenz als Kriterium zur Beurteilung der Mächtigkeit aktuell am Markt befindlicher DSS-Werkzeuge herangezogen und deren Fehleranfälligkeit anhand einer zu entwickelnden Metrik gemessen werden. Hierzu werden ausgewählte Personen, die bereits über Erfahrung in der Modellierung betriebswirtschaftlicher Problemstellungen mittels DSS-Werkzeugen verfügen, die einzelnen Fälle mit einem Werkzeug implementieren. Danach wird die Güte der Modellierung anhand der Metrik überprüft.

Anschließend wird die Hypothese, dass eine konzeptuelle Modellierung die Fehleranfälligkeit eines MDSS reduzieren kann, untersucht. In diesem Zusammenhang werden aktuelle Konzepte und Ansätze auf die Möglichkeit untersucht, ob und inwiefern sie einen Beitrag bei der Umsetzung einer konzeptuellen Modellierungssprache für MDSS leisten können. Hierbei handelt es sich hauptsächlich um Konzepte multidimensionaler Modellierung, die ihren Ursprung im On-Line Analytical Processing haben, wie z.B. die in Kapitel 3 erwähnten Ansätze SDWM und MML.

Nach der Evaluation bestehender Ansätze wird aufgrund der gewonnenen Erkenntnisse ein Vorschlag einer Modellierungssprache für modellgetriebene DSS erarbeitet. Die einzelnen Differenzierungsmerkmale der Fallklassifizierung aus Teil 1 werden durch Sprachkonstrukte und -attribute repräsentiert. Für die technische Realisierung ist die eXtensible Markup Language (XML) vorgesehen.

Abschließend könnte eine prototypische Implementierung der konstruierten Modellierungssprache entwickelt und dokumentiert werden. Neben der technischen Realisierung sollen in einer abschließenden Evaluation Relevanz, Möglichkeiten und Grenzen der konzeptuellen Modellierungssprache ausgearbeitet werden, indem ein Test anhand der erarbeiteten Referenz durchgeführt wird. Wiederum sollen die ausgewählten Personen die einzelnen Fälle implementieren und anschließend die Güte der Modellierung überprüfen, jedoch ausgehend von Spezifikationen in der konzeptuellen Modellierungssprache. Für den Vergleich und die Darstellung der Ergebnisse wird wiederum die bereits ausgearbeitete Metrik herangezogen.

5 Literatur

Alter, S. (1996). *Information Systems: A management perspective*. Menlo Park, CA: Benjamin / Cummings Publishing.

Böhnlein, A., Ende, U.-v. (2001). Ein konzeptuelles Data Warehouse-Modell für die Erstellung multidimensionaler Datenstrukturen. In Gesellschaft für Informatik, Fachgruppe 5.10 Informationssystem-Architekturen (Hrsg.), *Rundbrief 8. Jahrgang* (S. 25-57).

Codd, E., Codd, S., & Salley, C. (1993). *Providing OLAP to User-Analysts: An IT Mandate*. Abgerufen von: http://www.minet.uni-jena.de/dbis/lehre/ss2005/sem_dwh/lit/Cod93.pdf (am 15.06.2009).

Coles, S., & Rowley, J. (1996). Spreadsheet modelling for management decision making. *Industrial Management & Data Systems*, 96(7), 17-23.

Frank, U. (2007). Ein Vorschlag zur Konfiguration von Forschungsmethoden in der Wirtschaftsinformatik, in: F. Lehner; S. Zelewski (Hrsg.), *Wissenschaftstheoretische Fundierung und wissenschaftliche Orientierung der Wirtschaftsinformatik* (S.155-184). Berlin, Deutschland: Gito-Verlag.

Green, P. (1997). Use of Information Systems Analysis and Design (ISAD) Grammars in Combination in Upper CASE Tools – An Ontological Evaluation. In K. Siau, Y. Wand, & J. Parsons (Hrsg.), Proceedings of the 2nd CAI-SE/IFIP8.1 International Workshop on the Evaluation of Modeling Methods in Systems Analysis and Design (EMM-SAD'97) (S. 1-12). Barcelona, Spanien.

Herden, O., Harren, A. (2004). Die ODAWA-Methodik für den Entwurf von Data-Warehouse-Datenbanken. *Informatik - Forschung und Entwicklung*, 19(2), 87-96.

Holsapple, C. W., & Whinston, A. B. (1996). *Decision Support Systems: a knowledge-based approach*. Minneapolis, MN: West Publishing.

Holsapple, C. W. (2008). DSS Architecture and Types. In F. Burstein, & C. W. Holsapple (Hrsg.), *Handbook on Decision Support Systems 1* (S. 163-189). Berlin, Deutschland: Springer-Verlag.

Insua, D. R. (1990). *Sensitivity analysis in multi-objective decision making* (Lecture Notes in Economic and Mathematical Systems). Berlin, Deutschland: Springer-Verlag.

Koutsoukis, N. S., Mitra, G., & Lucas, C. (1999). Adapting on-line analytical processing for decision modelling: the interaction of information and decision technologies. *Decision Support Systems*, 26(1), 1-30.

March, S. T., & Hevner, A. R. (2007). Integrated decision support systems: A data warehousing perspective. *Decision Support Systems*, 43(3), 1031-1043

Mayer, M. K. (1998). Future trends in model management systems: parallel and distributed extensions. *Decision Support Systems*, 22(4), 325-335.

Panko, R., & Sprague, R. Jr. (1998). Hitting the wall: errors in developing and code inspecting a 'simple' spreadsheet model. *Decision Support Systems*, 22(4), 337-353.

Powell, S. G., Baker, K. R., & Lawson, B. (2008). A critical review of the literature on spreadsheet errors. *Decision Support Systems*, 46(1), 128-138.

Power, D., & Sharda, R. (2007). Model-driven decision support systems: Concepts and research directions. *Decision Support Systems*, 43, 1044-1061.

Rieger, B. (1993). Der rechnergestützte Arbeitsplatz für Führungskräfte. Habilitationsschrift, Technische Universität Berlin.

Simon, H. A. (1977). *The new science of management decision*. Englewood Cliffs, NJ: Prentice-Hall.

Sprague, R. Jr., & Watson, H. (1996). *Decision Support for Management*. Upper Saddle River, NJ: Prentice-Hall.

Turban, E., Aronson, J., Liang, T.-P., & Sharda, R. (2007). *Decision Support and Business Intelligence Systems*. Upper Saddle River, NJ: Pearson.

Wand, Y., & Weber, R. (2002). Research Commentary: Information Systems and Conceptual Modeling – A Research Agenda. *Information Systems Research*, 13(4): 363-376.

Werners, B. (2008). *Grundlagen des Operations Research*. Berlin, Deutschland: Springer-Verlag.