

Wissensbasiertes Business Intelligence für die Informations-Selbstversorgung von Entscheidungsträgern

Matthias Mertens
OFFIS - Institut für Informatik
Escherweg 2
26121 Oldenburg, Germany
mertens@offis.de

ABSTRACT

Im Bereich der Business Intelligence haben sich Analytische Informationssysteme (AIS) mit dem Ziel entwickelt, verschiedene Datenquellen integriert analysieren zu können und Informationen zu gewinnen die Business User, in ihrem Entscheidungsfindungsprozess unterstützen. Sowohl die hohe Komplexität, die sich aus der Flexibilität und Mächtigkeit solcher Systeme ergibt, als auch die hohe notwendige Interaktion mit dem Nutzer zur Durchführung adäquater Analysen, bedingen entsprechendes Analyse- und Domänenwissen sowie ein tiefergehendes konzeptionelles und technisches Verständnis. Dieses ist meistens bei Business Usern ohne entsprechende Schulung nicht gegeben, wodurch eine eigenständige Informationsversorgung mittels AIS behindert wird. Erschwerend kommt hinzu, dass AIS in der Regel keine zusätzlichen Metadaten zu Business Regeln, Strategien oder Hintergrundinformationen erfassen, verwalten und für die Analyseunterstützung bereitstellen können.

Idealerweise sollten Business User auf Basis einer Analyseunterstützung des AIS dazu befähigt werden, adäquate Analysen durchzuführen, ohne zwingend über Analyse- und Domänenwissen verfügen zu müssen. Diese analyseunterstützenden Funktionalitäten können von weiterführenden Informationen, über eine Navigationsunterstützung für Analysepfade bis hin zu einer Vorschlagsgenerierung von Analyse-schritten reichen.

In diesem Paper werden Konzepte eines Analyseprozesses und darauf aufbauend analyseunterstützende Funktionen vorgestellt, die eine Informations-Selbstversorgung des Business Users erlauben. Der Fokus wird hierbei auf die Erweiterung eines AIS um semantische Metadaten gelegt, um eine Erfassung, Verwaltung und Nutzung von Analyse- und Domänenwissen zu ermöglichen.

Categories and Subject Descriptors

J.1 [Administrative data processing]: Business; H.4.2 [Information Systems Applications]: Types of Systems—*Decision support*

23rd GI-Workshop on Foundations of Databases (Grundlagen von Datenbanken), 31.05.2011 - 03.06.2011, Obergurgl, Austria.
Copyright is held by the author/owner(s).

Keywords

Semantic Metadata, Data Warehouse, Analytical Information System, Decision support, Business Intelligence

1. EINLEITUNG

Die Unterstützung des Managements und die Verbesserung des Entscheidungsfindungsprozesses werden als Schlüsseleigenschaften der Business Intelligence (BI) gesehen [6]. Entscheidungsträger, sogenannte Business User, sollen befähigt werden, alle benötigten Informationen zur richtigen Zeit zu erhalten. In der BI wurden analytische Informationssysteme (AIS) entwickelt, die es Business Usern erlauben große Datenmengen zu visualisieren, zu handhaben und zu analysieren. AIS bestehen aus einem Data Warehouse (DWH) und darauf aufbauenden Analysekomponenten. Während das DWH es ermöglicht verschiedene Datenquellen qualitätsgesichert zu integrieren und multidimensional aufzubereiten, erlauben die Analysekomponenten Online Analytical Processing (OLAP) Operatoren, komplexe statistische Verfahren sowie geografische Operatoren in verschiedenen Visualisierungen auf den integrierten Daten durchzuführen.

AIS haben jedoch auch verschiedene Mängel, die im nächsten Abschnitt 2 näher betrachtet werden. Aus diesen leitet sich die Forschungsfrage sowie die zugehörigen Anforderungen an einen Ansatz ab, welcher in Abschnitt 3 diskutiert wird. Im Anschluss wird der eigene Ansatz mit zugehörigen Konzepten und Funktionalitäten zur Analyseunterstützung in Abschnitt 4 mit der Anwendungsdomäne „Krankenhausmarktanalyse“ (KMA) präsentiert, bevor im Abschnitt 5 verwandte Arbeiten in diesem Forschungsumfeld aufgezeigt werden. Abschließend erfolgt in Abschnitt 6 eine Zusammenfassung sowie ein Ausblick des Forschungsvorhabens.

2. MÄNGEL ANALYTISCHER INFORMATIONSSYSTEME

In dem am OFFIS entwickelten AIS - Multidimensional Statistical Data Analysis Engine (MUSTANG) [9] - konnte beobachtet werden, dass sich aus der hohen Flexibilität und Mächtigkeit von AIS eine Komplexität ergibt, welche zu einer signifikanten Herausforderung für Business User werden kann, wenn diese eigenständig adäquate explorative Analysen durchführen möchten. Im Gegensatz zu Analyseten verfügen Business User in der Regel über ein geringeres konzeptionelles Verständnis des multidimensionalen Datenmodells (MDM) sowie geringeres notwendiges Analyse- und

Domänenwissen. Hierbei enthält das Domänenwissen Informationen darüber, welche Fragestellungen einer zu untersuchenden Analyse zu Grunde liegen, mit welchen Kennzahlen Analysen zu spezifischen Fragestellungen möglich sind, welche explizite Semantik diese Kennzahlen haben und in welchen Beziehungen diese zueinander stehen. In Abgrenzung dazu umfasst Analysewissen Informationen zu den Analyseinstrumenten, d.h den Analyseoperationen, -verfahren und den möglichen Visualisierungen. Schließlich wird in den AIS für die Durchführung komplexer Analysen eine hohe Interaktion mit den Business Usern benötigt, wodurch diese ohne entsprechende Schulung schnell überfordert sind.

Ein weiterer Mangel von AIS wird in der geringen Berücksichtigung von Metadaten gesehen [2], die weiterführende Informationen über die quantitativen DWH Daten sowie die DWH Struktur (Kennzahlen und Dimensionselemente) bereitstellen. Zu diesen zählen Annahmen, Definitionen, Business Regeln, Terminologien und Hintergrundinformationen [10]. Daher müssen Business User sich die Semantiken von Daten und Strukturen durch Zuhilfenahme externer Quellen selbst erschließen [2].

Des Weiteren sind AIS derzeit nicht in der Lage, Analyse- und Domänenwissen zu importieren, zu verwalten und für weiterführende Analyseunterstützung der Business User zu nutzen. Insbesondere ist die Expertise aus dem Bereich der Analysestrategien und den „Best Practices“ für Analysen in spezifischen Domänen von Interesse. Wissen, das durch Analysten in AIS eingebracht wird, geht in der Regel verloren [1].

Idealerweise sollten Business User befähigt werden, multidimensionale quantitative Daten zu analysieren, ohne zwingend über Analyse- und Domänenwissen sowie ein tiefgehendes konzeptionelles Verständnis verfügen zu müssen. Das AIS sollte den Business User in seinen explorativen ad-hoc Analysen unterstützen, indem es modelliertes semantisches, maschinenlesbares und -verständliches Wissen ausnutzt. Analyseunterstützende Funktionalitäten für eine Business User Self-Information Service [2, 12], könnten von weiterführenden Informationen zu DWH Entitäten, über eine Navigationsunterstützung für Analysepfade bis hin zu einer Vorschlagsgenerierung von weiteren sinnvollen Analyseschritten reichen.

In diesem Paper werden nun im Folgenden Konzepte für eine semantische Metadatenbasis sowie darauf aufbauende analyseunterstützende Funktionalitäten für ein AIS vorgestellt. Ziel ist es, dass Business User sich selbst mit Informationen auf eine effiziente und intuitive Art und Weise versorgen können. Exemplarisch werden Beispiele aus der Domäne Krankenhausmarktanalyse (KMA) gebracht, da hier Business User, z.B. Krankenhauscontroller durch die Veränderungen im deutschen Gesundheitswesen gezwungen werden, zur Sicherung der Wettbewerbsfähigkeit die Potentiale ihres Krankenhauses zielgerichtet zu erschließen und Leistungsangebote konkurrenzfähig auszurichten.

3. FORSCHUNGSFRAGE UND ANFORDERUNGEN

Motiviert aus den in Abschnitt 2 genannten Mängeln wird die folgende Fragestellung abgeleitet:

Wie kann eine Komplexitätsreduktion von AIS mit dem Ziel erfolgen, ungelernete Business User zu befähigen, selbst adäquate explorative Analysen für eine intuitive und effiziente Informations-Selbstversorgung durchzuführen.

Um die Forschungsfrage zu beantworten, sieht der verfolgte Lösungsansatz die Verwendung von verschiedenen Analyseunterstützungsfunktionalitäten vor, welche explizites modelliertes semantisches Domänen- und Analysewissen ausnutzen. Dieses wird in ein semantisches Datenmodell des AIS importiert, dort verwaltet und durch verschiedene Services genutzt werden.

Folgenden Anforderungen werden an einen Ansatz gestellt:

- Der Ansatz soll es einem domänenunabhängigen AIS ermöglichen Business User im Kontext domänenspezifischer Analyseaufgaben zu unterstützen. Dafür wird ein Konzept bestehend aus einem generischen Datenmodell in Verbindung mit einem domänenunabhängigen AIS benötigt. Das Konzept wird für eine spezifische Domäne instanziiert.
- Der Ansatz soll den Lernprozess von Analyse- und Domänenwissen unterstützen und den initialen Einarbeitungs- und Trainingsaufwand ins AIS reduzieren. Business User sollen von der inhärenten Expertise des AIS lernen.
- Der Ansatz soll eine Komplexitätsreduktion eines AIS verfolgen, ohne jedoch die Flexibilität und die Mächtigkeit des AIS einzuschränken. Das AIS soll für Business User mit geringer Expertise intuitiv benutzbar sein, so dass diese adäquate explorative Analysen auf dem DWH durchführen können. Des Weiteren soll der Ansatz die Anzahl der Nutzer-Interaktionen für das Erreichen gleicher Analyseergebnisse reduzieren, wodurch die Effektivität des AIS gesteigert werden soll.
- Das Konzept soll es ermöglichen, mit verschiedenen Metadatenarten umzugehen und diese in einer intelligenten Art und Weise für eine Analyseunterstützung zu verknüpfen. Oftmals sind Ansätze in der Literatur zu finden [2, 7, 8, 11], die Hintergrundinformationen sowie Regeln über die DWH Struktur bereithalten. Ferner sollen aber Metadaten zum Analyseprozess berücksichtigt werden, die den Analyseprozess von einer Fragestellung über verschiedene Analyseschritte hin zu einem Analyseergebnis beschreiben und wesentliche Informationen für eine Analyseunterstützung bereithalten. Diese Metadaten werden auch Strategien oder „Best Practices“ genannt und beschreiben die Expertise bzw. das Analyse- und Domänenwissen eines Analysten. Schließlich soll das AIS ebenfalls Metadaten zu quantitativen Daten des DWH, sprich zu Analyseergebnissen wie z.B. Trends, spezifische Zusammenhänge, etc. verarbeiten.

Im nächsten Abschnitt 4 werden verschiedene Konzepte eines Datenmodells sowie die darauf anwendbaren Funktionalitäten zur Analyseunterstützung und zur Erfüllung der Anforderungen diskutiert.

4. WISSENSBASIERTE FUNKTIONEN UND KONZEPTE FÜR EINE ANALYSEUNTERSTÜTZUNG

Ziel des hier diskutierten Ansatzes ist es, Business User im Kontext domänenspezifischer Analyseaufgaben zu unterstützen. Hierzu sollen verschiedene unterstützende Funktionalitäten auf Basis eines modellierten semantischen Metadatenmodells bereitgestellt werden (s. Abschnitt 4.1). Als Wissensrepräsentationssprache zur Modellierung des Metadatenmodells, in Form von mehreren miteinander verknüpften Ontologien, kommt OWL-DL zum Einsatz. Eine DWH-Ontologie bildet das MDM des jeweils zu Grunde liegenden DWHs als Instanzen ab und ermöglicht es, in den darauf aufbauenden Ontologien das MDM zu referenzieren. Die Analyse-Ontologie modelliert abstrakt Entitäten und deren Beziehungen, die die Analyseprozesse in einem MDM beschreiben. Diese decken den gesamten Analyseprozess von Fragestellungen, über Analyseketten hin zu Analyseergebnissen ab, beschreiben aber auch die enthaltenden Operatoren, Visualisierungen, Verfahren und Business Rules. Die wichtigsten Konzepte werden in Abschnitt 4.2 erläutert. Die auf der Analyse-Ontologie aufbauende Domänen-Ontologie beinhaltet die konkreten Instanzen der Analyse-Ontologie für eine Domäne. Sie bildet also das für Analyseunterstützung genutzte Analyse- und Domänenwissen ab. Für eine detaillierte Beschreibung der Ontologien und deren Zusammenhänge sei auf [9] verwiesen.

Im Metadatenmodell zu speicherndes Wissen soll Aussagen zu konkreten Fragestellungen, über durchzuführende Analyseschritte bis hin zu erkennbaren Analyseergebnissen enthalten. Soll z.B. die Fragestellung „Welchen Marktanteil hat mein Krankenhaus (KH) im Einzugsgebiet?“ untersucht werden, so ist als Wissen modelliert, dass das Einzugsgebiet, bestehend aus Kernmarkt, erweiterter Kernmarkt und Peripheriemarkt, für das KH über dessen Fallzahlen ermittelt werden muss und dass der Marktanteil eine berechnete Kennzahl im DWH ist. Diese berechnet sich aus dem Verhältnis der erwarteten Fallzahl und den behandelten Fällen des KH. Des Weiteren ist z.B. als Wissen modelliert, dass sich aus der ersten Fragestellung weitere relevante Fragestellungen ableiten können, z.B. „Wie verhält sich die Fragestellung für Konkurrenten oder aber für spezifische Fachabteilungen (FA) des eigenen KH?“. Für letzteres müssen insbesondere die behandelten Diagnosen (ICD-Codes) und die vorhandene Ausstattung / Verfahren (OPS-Codes) sowie der Versorgungsschwerpunkt der FA im MDM berücksichtigt werden. Besonders relevant ist das Wissen zu weiterem Analyseverfahren im Kontext von Analyseergebnissen, wie z.B. erkannten Auffälligkeiten: Änderungen im Patientenspektrum / Einweiserverhalten; oder Erkennen von Versorgungslücken oder Regionen mit „stillen Reserven“.

Das Konzept sieht einen initialen Aufbau der Wissensbasis, mit Hilfe einer Expertengruppe vor. Diese können im KMA Umfeld z.B. einer Krankenhauskette oder einem Beratungsunternehmen angehören und ihr Analyse- und Domänenwissen in den Ontologien persistieren. Da das modellierte Wissen eine Allgemeingültigkeit in der modellierten Domäne haben soll, ist eine spätere personalisierte Anpassung bzw. die Erfassung zusätzlicher personenbezogener Metadaten derzeit nicht vorgesehen.

4.1 Analyseunterstützende Funktionalitäten

Um den genannten Mängeln aus Abschnitt 2 unter Berücksichtigung der in Abschnitt 3 erläuterten Anforderungen zu begegnen, werden die Metadaten in Form einer semantischen Suche, Navigation und eines Recommendations genutzt. Die Metadaten werden an den Business User über diese Funktionalitäten kommuniziert.

Semantische Suche: Das übergeordnete Ziel der semantischen Suche ist, eine Suchfunktion auf definierten Metadaten anzubieten. Instanzen des Metadatenmodells sollen anhand ihrer Semantik gesucht, gefunden und anschließend visualisiert werden.

Ein spezifischeres Ziel ist unter anderem die Suche nach bestehenden Fragestellungen und damit verbundenen Analyseketten, die sich auf die jeweilige Analysesituation adaptieren lassen. Auch semantisch verknüpfte Fragestellungen, die in einer Vorgänger-, Nachfolgerbeziehung stehen, können gefunden werden. Daneben spielt auch das Finden von Analyseergebnissen und damit verbundenen quantitativen Daten des DWH eine Rolle. Annotationen auf den quantitativen Daten können mit ihrer Semantik gefunden und in der zugehörigen Analysevisualisierung wieder dargestellt werden. Ebenfalls lassen sich die durchlaufende Analyseketten samt Fragestellung und analysierendem Business User ermitteln und als Ausgangsbasis für anknüpfende Analysen verwenden.

Semantische Navigation: Werden die Klassen und Relationen des Metadatenmodells für eine konkrete Domäne instanziiert, so kann eine Navigation von einer Klasseninstanz zur Nächsten, entlang der dazwischen definierten semantischen Relation, erfolgen. Im Kontext eines AIS kann diese semantische Navigation zum einen als eine reine Navigation innerhalb der Metadaten erfolgen und zum anderen können zusätzliche assoziierte Operationen des AIS ausgeführt werden.

Für Ersteres sei auf Abb. 1 verwiesen. Wird die Fragestellung zur Kennzahl B mit Hilfe der semantischen Suche gefunden, so kann der Business User sowohl zu den notwendigen Vorgänger-Fragestellungen als auch den Nachfolger-Fragestellungen oder zu weiteren verbundenen Instanzen navigieren.

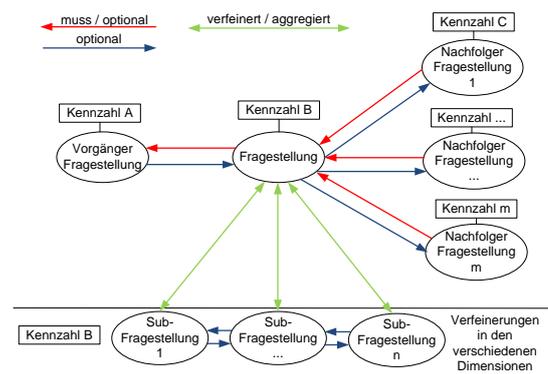


Abbildung 1: Beziehungen von Fragestellungen

Möchte der Business User basierend auf einer Fragestellung eine konkrete Analyse durchführen, so kann er von der

Fragestellung zu der verbundenen Start-Analysevisualisierung wechseln und dann entlang eines spezifischen Analysepfades zu einer Ende-Analysevisualisierung navigieren (s. Abb. 2). Diese Art der semantischen Navigation ist immer mit realen Aktionen und Daten des AIS verbunden, da Analysevisualisierungen quantitative Daten des DWH anzeigen und beim Analysevisualisierungswechsel durch die Domänenoperationen ein oder mehrere OLAP Operationen im AIS angewendet werden.

Vorschlagsgenerierung: Eine wichtige Funktion ist das Geben von Hinweisen und Vorschlägen durch das AIS im Kontext einer Fragestellung und einer Analysevisualisierung. Diese werden aus dem im Metadatenmodell hinterlegten Analyse- und Domänenwissen abgeleitet. Vor allem weiterführende Hintergrundinformationen, Business Rules und Analysestrategien sowie weitere Fragestellungen können wertvolle Informationen im Analyseprozess sein.

Hinweise zu weiterführenden sinnvollen Analysevisualisierungen können durch das AIS gegeben werden, indem mögliche Domänenoperationen und die enthaltenen Business Rules auf ihre Anwendbarkeit in einer Analysevisualisierung überprüft werden. Ziel der Vorschlagsgenerierung soll es sein, dass relevante Analysevisualisierungen erreicht werden, die eine Interpretation der quantitativen Daten hinsichtlich der Fragestellung zulassen. Wichtig ist, dass beim Geben der Hinweise und Vorschläge sowie bei der Anwendung von Domänenoperationen die Hintergründe kurz erläutert werden, damit diese für den Business User nachvollziehbar bleiben und er sich zusätzlich Analyse- und Domänenwissen aneignen kann.

4.2 Wissensbasierte Konzepte

Im Folgenden werden die für eine Analyseunterstützung notwendigen Entitäten des Analyseprozesses in einem AIS näher betrachtet.

Fragestellung: Eine Fragestellung gibt einer Hypothese Ausdruck, die es zu klären gilt und auf deren Ergebnis eine Entscheidung basieren kann. Als Beispiel sei "Was sind die Marktanteile in meinem Einzugsgebiet?" genannt. Hierbei ist eine einzelne Fragestellung in der Regel nicht losgelöst von Anderen zu betrachten (s. Abb. 1). Vielmehr kann eine Fragestellung über Vorgänger verfügen, wie z.B. "Was ist mein Einzugsgebiet?", auf welche zuvor eine Antwort gefunden werden muss. Sie kann über weiterführende Nachfolger-Fragestellungen verfügen, die sich aus dem Analyseergebnis ergeben und optional weiter analysiert werden können. Beispiele sind: "Was sind die Marktpotenziale in meinem Einzugsgebiet?" oder "Wie wird sich mein Marktanteil entwickeln?". Neben Vorgängern und Nachfolgern kann es auch Subfragestellungen geben, welche die gleiche Kennzahl(en) wie die Originalfrage behandeln, jedoch diese hinsichtlich ein oder mehrerer Dimensionen weiter verfeinern.

Analysekette: Eine Analysekette wird für die Analyse von Fragestellungen verwendet. Diese bildet den logischen Rahmen der Analyse und enthält weitere Konzepte wie Analysepfade, Analysevisualisierungen, Domänenoperationen und Business Rules. Die Zusammenhänge sind in der Abb. 2 dargestellt. Ziel der Analysekette ist das Finden ein oder mehrerer Analysevisualisierungen, die eine Interpretation der quantitativen Daten des DWH erlauben.

Analysevisualisierung: Eine Analysevisualisierung ist die grafische Repräsentation der quantitativen Daten eines Analyseschrittes z.B. in Form einer Pivottabelle, eines Diagramms oder einer Karte. Sie besteht aus 1 bis n Kennzahlen, die wiederum in 1 bis m Dimensionen aufgespannt sind. Von diesen Dimensionen sind jeweils 1 bis z Dimensionselemente gewählt. Im Kontext der Analysekette sind sogenannte Start- und Ende-Analysevisualisierungen definiert. Erstere dienen als Einstiegspunkte für die Analysen von Fragestellungen. Von ihnen aus können die Analysepfade mit ihren verschiedenen Analysevisualisierungen durchlaufen werden. Letztere ermöglichen das Interpretieren der quantitativen Daten hinsichtlich der Fragestellung und das Ableiten von Analyseergebnissen.

Domänenoperation: Eine Domänenoperation ermöglicht das Navigieren zwischen zwei Analysevisualisierungen und kommt zum Einsatz, wenn spezifische Ausprägungen des MDM in einer Analysevisualisierung eintreten. Domänenoperationen können aus einer Menge auszuführender OLAP-Operatoren, Business Rules und Visualisierungswechseln bestehen. Ihr Zweck ist es, die notwendigen Schritte der Kennzahlen-, Dimensionen-, Dimensionselemente- und Visualisierungsauswahl bzw. des -wechsels für den Endanwender durchzuführen und somit die Komplexität des MDM und der Analysedurchführung zu verbergen. So kann der Business User direkt von einer Analysevisualisierung zu einer nächsten sinnvollen Analysevisualisierung gelangen. Ob eine Analysevisualisierung als sinnvoller weiterer Analyseschritt gesehen werden kann, ist über die Business Rules definiert.

Business Rules: Business Rules repräsentieren konkretes Analyse- und Domänenwissen und werden als semantische Metadaten zu den domänenspezifischen Inhalten eines AIS modelliert. Sie beschreiben weiterführende Informationen und Regeln und kommen im Analyseprozess in den Domänenoperationen zur Anwendung. Business Rules lassen sich aus den „Best Practices“ eines Analysten ableiten und beziehen sich in der Regel auf eine Fragestellung, eine Analysevisualisierung oder ein Analyseergebnis, oder aber auf eine Kombination aus diesen. Die Business Rule ist anwendbar, wenn das MDM einen spezifischen definierten Zustand erreicht. Mehrere Business Rules können in unterschiedlichen Domänenoperationen anwendbar sein, woraus sich mehrere Möglichkeiten für den Business User ergeben können, wie er seine Analyse fortsetzen möchte. Durch die Gewichtung der Business Rules kann hierbei ein Ranking entstehen.

Analysepfad: Unter einem Analysepfad wird eine Menge von Analysevisualisierungen verstanden, die durch Domänenoperationen zu einem Pfad in einer definierten Reihenfolge verbunden werden. Typischerweise wird ein Analysepfad in einer konkreten Analyse von einer Start-Analysevisualisierung zu einer Ende-Analysevisualisierung durchlaufen. Hierbei ist zu beachten, dass Analysepfade nicht zwingend zuvor definiert sein müssen, sondern sich aus der Anwendbarkeit von Domänenoperationen auf Analysevisualisierungen im Kontext einer spezifischen Fragestellung ergeben können. Die Menge aller Analysepfade zu einer Fragestellung bildet die Analysekette.

Analyseergebnis: Analyseergebnisse beschreiben eine Interpretation von quantitativen Daten im MDM des DWH.

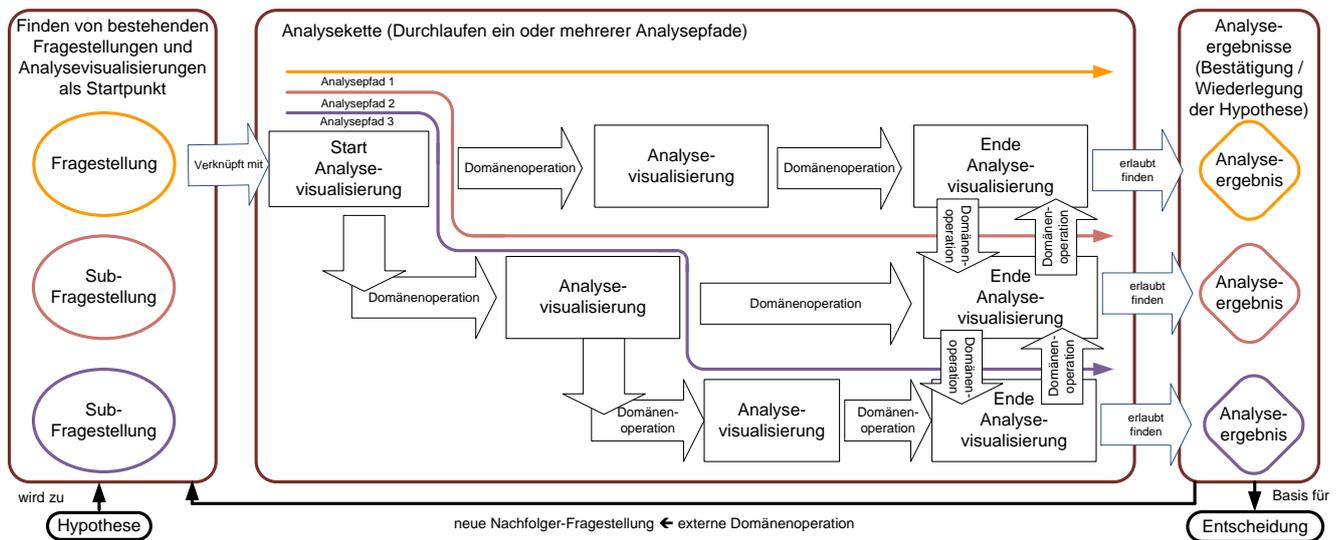


Abbildung 2: Zusammenhang von Konzepten im Analyseprozess

Dies können Korrelationen und Zusammenhänge zwischen den Daten, aber auch Trends, Einbrüche, Zunahmen, Auffälligkeiten, geografische Ballungen, etc. sein.

5. VERWANDTE ARBEITEN

Im Kontext der Business Intelligence gibt es vielfältige Forschungsfragen, die von der Erschließung des Analysewissens [1], über die Anpassung der BI-Tools an Business Anforderungen [11, 2] und die integrierte Anfrage von BI-Tools [12, 4] bis hin zur Annotation von Datenschemata mit weiterführenden Informationen zur Analyseunterstützung reichen [2, 7, 8, 11]. In vielen Fällen soll der Analyst und teilweise auch der Business User [2, 12] stärker in den Fokus rücken, indem die BI-Software besser an individuelle Informationsbedürfnisse angepasst werden kann oder aber eine Analyseunterstützung zur individuellen Befriedigung des Informationsbedarfs angeboten wird. Viele der verwandten Arbeiten versuchen die jeweiligen adressierten Probleme durch die Nutzung von semantischen Metadaten und damit verbundenen Semantic Web Technologien zu lösen.

Im Bereich der Dokumentation von Analyseprozessen und -ergebnissen ist die Arbeit [1] zur Distribution von Business-Intelligence-Wissen zu nennen. Diese zielt auf die kontrollierte organisationsweite Verbreitung und Weiterverwendung von Berichten und Analyseansätzen durch das Einstellen der BI-Inhalte in Wissensmanagementsysteme ab, jedoch ohne auf eine konkrete technische Umsetzung einzugehen.

Im Kontext von Analyseprozessen ist auch die Arbeit von [3] relevant, da diese den Begriff der Analyseketten einführt. Fokus der Arbeit ist die Mensch-Maschine Interaktion, wobei das Finden, Aufbereiten und Darstellen von Daten als technisch / operative Tätigkeit gesehen wird. Zu den kognitiven Tätigkeiten zählen die Bildung eines anwenderorientierten Analysemodells sowie dessen Prüfung und Verfeinerung.

Während der Durchführung von Analysen werden eine Reihe von OLAP Anfragen gesendet, um durch den mul-

tidimensionalen Datenraum zu navigieren und um die benötigten Informationen zu erfragen. Hierbei ist laut [5] das Erstellen entsprechender Anfragen eine schwierige Aufgabe. Daher wird in dieser Arbeit ein Framework vorgestellt, das den Nutzer in der Analyse unterstützt, indem es passende OLAP Operatoren vorschlägt. Diese werden durch die Auswertung des OLAP Server Query Logs abgeleitet.

Eine andere Art der Analyseunterstützung für eine höhere Nutzerfreundlichkeit wird im Kontext der BI in der Trennung von Business und IT Belangen gesehen. Anders als in den zuvor genannten Arbeiten spielen hier semantische Metadaten eine entscheidende Rolle. Über die Metadatenmodelle können die Entitäten der zugrunde liegenden BI-Systeme mit ihren zugehörigen Relationen modelliert werden. Ebenfalls erlauben diese flexiblen, erweiterbaren Metadatenstrukturen weiterführende Informationen zu den Entitäten, wie beispielsweise Business Rules oder Expertenwissen festzuhalten. In [7] und [8] werden diese Möglichkeiten weiter vertieft und als Anwendung z.B. die Nutzung der Metadaten im Extraktions-, Transformations- und Ladeprozess (ETL) für ein DWH angeführt.

In der Arbeit von [11] wird auf dieser Semantic Web Basis eine Architektur für analytische Tools vorgeschlagen, um Effizienzsteigerungen im Entscheidungsfindungsprozess zu erzielen. Durch die Nutzung einer Domänen-Ontologie, welche die Entitäten der zu untersuchenden Domäne, deren Relationen sowie weitere Informationen bereitstellt und durch die Nutzung einer Business-Intelligence-Ontologie, welche Informationen zu den Datenstrukturen des multidimensionalen Modells vorhält, soll eine Unterstützung im Analyseprozess erfolgen. Zum einen soll die Auswahl von Dimensionselementen durch proaktive Vorschläge erleichtert werden und zum anderen sollen durch ein semiautomatisches Umschreiben der Anfragen Analyseergebnisse aufgewertet werden. Die Entitäten werden in den jeweiligen Domänen-Ontologien oftmals mit ihrem natürlichsprachlichen Namen versehen, da diese eher die Business Semantik ausdrücken als die korrespondierenden technischen Bezeichner des MDM. Nutzer

können so Anfragen in einer für sie vertrauten Terminologie an das System stellen. Dieser Ansatz wird in [4] und [12] verwendet, um auf einer abstrakten Ebene unternehmensweite Informationen integriert aus verschiedenen BI-Systemen wie DWH, ERP, CRM, etc. anzufragen. In diesen Ansätzen werden semantische Metadaten auch für die Integrationsaufgaben und das Umschreiben von Anfragen verwendet.

Die Unterstützung eines technisch unversierten Business Users wird insbesondere in [2] und [12] fokussiert. In [2] liegt der Schwerpunkt auf einer kollaborativen ad-hoc Entscheidungsunterstützung, in der Daten integriert aus den verschiedenen BI-Tools dargestellt und über semantische Metadaten mit weiterführenden Informationen versehen werden. Insbesondere sollen über Web 2.0 Technologien Information Mash-Ups gebildet, aber auch eine Kollaboration zwischen verschiedenen Business Usern erzielt werden.

Auch wenn, wie in vielen anderen Arbeiten, die semantischen Metadaten eine entscheidende Rolle spielen, grenzt der eigene Ansatz sich von diesen durch die Modellierung von Konzepten und deren Instanziierung in Form von Analyse- und Domänenwissen ab. Im Fokus steht dabei das Wissen zu Analyseprozessen, das von verschiedenen Komponenten eines AIS zur Unterstützung des Business Users genutzt werden kann. Eine technologisch ähnliche Umsetzung mit Hilfe von verschiedenen Ontologien wie sie in [11] und [12] genannt werden, wird angestrebt.

6. FAZIT UND AUSBLICK

Analytische Informationssysteme haben sich im Kontext der Business Intelligence als Systeme zur Informationsgewinnung für Business User im Entscheidungsfindungsprozess etabliert. Allerdings setzen AIS aufgrund ihrer hohen Interaktionsmöglichkeiten und Komplexität entsprechendes Analyse- und Domänenwissen sowie ein tiefergehendes technisches Verständnis voraus, um adäquate Analysen durchführen zu können (s. Abschnitt 1). Dieses liegt jedoch ohne entsprechende Schulungen bei Business Usern nicht vor. Auch werden in der Regel keine Metadaten zu Business Regeln, Strategien oder Hintergrundinformationen durch AIS bereitgestellt, die den Business User unterstützen können. Da jedoch Business User befähigt werden sollen, sich selbst mit Informationen zu versorgen, wurde in dieser Arbeit als Forschungsfrage untersucht, wie die Komplexität von AIS reduziert werden kann, damit Business User eine adäquate explorative ad-hoc Analyse durchführen können. Hierzu wurden zunächst in Abschnitt 3 Anforderungen an einen entsprechenden Ansatz definiert. Darauf aufbauend wurden in Abschnitt 4 verschiedene Konzepte aus dem Bereich der Datenanalyse definiert und erläutert, die in einem semantischen Metadatenmodell für ein AIS modelliert und instanziiert werden können und somit als Basis für weiterführende Analyseunterstützungsfunktionalitäten dienen. Von zentraler Bedeutung waren die Konzepte der Fragestellung, Analyseketten, Analysepfad und Analysevisualisierung sowie Domänenoperationen, Business Rules und Analyseergebnisse. Als Funktionen wurden die Semantische Suche, die Semantische Navigation und die Vorschlagsgenerierung präsentiert, bevor in Abschnitt 5 verwandte Arbeiten vorgestellt und gegen den eigenen Ansatz abgegrenzt wurden. Als weiterer Schritt im Forschungsvorhaben werden die vorgestellten Konzepte in einem Metadatenmodell mittels Wissensrepräsentations-

prachen modelliert und für die Domäne der Krankenhausmarktanalyse instanziiert. Dieses Metadatenmodell wird in eine zu implementierende semantische Metadatenebene des am OFFIS entwickelten AIS - Multidimensional Statistical Data Analysis Engine (MUSTANG) - eingebettet, um Analyse- und Domänenwissen zu erfassen, zu verwalten und für die genannten Analyseunterstützungsfunktionalitäten zu verwenden. Eine Umsetzung und Evaluierung wird im Rahmen einer laufenden Dissertation und einer studentischen Projektgruppe erfolgen.

7. REFERENCES

- [1] H. Baars. Distribution von Business-Intelligence-Wissen. *Analytische Informationssysteme*, pages 409–424, 2005.
- [2] H. Berthold, P. Rösch, S. Zöller, F. Wortmann, A. Carenini, S. Campbell, P. Bisson, and F. Strohmaier. An Architecture for ad-hoc and collaborative Business Intelligence. In *Proceedings of the 2010 EDBT/ICDT Workshops*, EDBT '10, pages 13:1–13:6, New York, NY, USA, 2010. ACM.
- [3] N. Bissantz. Deltaminer. *Wirtschaftsinformatik*, 43(1):77–80, 2001.
- [4] L. Cao, C. Zhang, and J. Liu. Ontology-based Integration of Business Intelligence. *Web Intelligence and Agent Systems*, Volume 4, 2006.
- [5] A. Giacometti, P. Marcel, and E. Negre. A Framework for Recommending Olap Queries. In *Proceeding of the ACM 11th international workshop on Data warehousing and OLAP*, DOLAP '08, pages 73–80, New York, NY, USA, 2008. ACM.
- [6] P. Gluchowski and H.-G. Kemper. Quo Vadis Business Intelligence? Aktuelle Konzepte und Entwicklungstrends. *Business Intelligence Spektrum*, 1. Jg., Heft 1, Mai 2006:12 – 19, 2006.
- [7] N. Inference. Ontology and Data Warehousing. Technology white paper, NETWORK INFERENCE, INC., 2004.
- [8] L. Ludwig. Business Intelligence und das Semantic web: Ein Traumpaar. 2005.
- [9] M. Mertens, Y. Teiken, and H.-J. Appelrath. Semantische Anreicherung von strukturierten Daten und Prozessen in Analytischen Informationssystemen am Beispiel von Mustang. In *Forschungskolloquium der GI Fachgruppe 5.8 - Management Support Systems, Dortmund, Deutschland*. Universität Dortmund, 2009.
- [10] B. O'Neil. Semantics and business. *The Data Administration*, 2007.
- [11] D. Sell, L. Cabral, E. Motta, J. Domingue, F. Hakimpour, and R. Pacheco. A semantic web based Architecture for Analytical Tools. In *CEC '05: Proceedings of the Seventh IEEE International Conference on E-Commerce Technology*, pages 347–354, Washington, DC, USA, 2005. IEEE Computer Society.
- [12] M. Spahn, J. Kleb, S. Grimm, and S. Scheidl. Supporting Business Intelligence by providing ontology-based End-User Information Self-Service. In *OBI '08: Proceedings of the first international workshop on Ontology-supported business intelligence*, pages 1–12, New York, NY, USA, 2008. ACM.