Bridging the Gap between Data Warehouses and Business Processes

A Business Intelligence Perspective for Event-Driven Process Chains

Veronika Stefanov[1] and Beate List[1]
Women's Postgraduate College for Internet Technologies
Institute of Software Technology and Interactive Systems
Vienna University of Technology
{stefanov, list}@wit.tuwien.ac.at

Josef Schiefer
Institute of Software Technology and Interactive Systems
Vienna University of Technology
js@ifs.tuwien.ac.at

Abstract

Data Warehouse (DWH) information is accessed by business processes, and sometimes may also initiate changes of the control flow of business process instances. Today, there are no conceptual models available that make the relationship between the DWH and the business processes transparent. In this paper, we extend the Event-Driven Process Chain, a business process modeling language, with an additional perspective to make this relationship explicit in a conceptual model. The model is tested with example business processes.

1. Introduction

A Data Warehouse (DWH) is more than just a big database. Defined as "a subject-oriented, integrated, time-variant, nonvolatile collection of data in support of management's decision-making process" [7], it plays a crucial role in modern organizations. More and more business processes depend on services provided by the DWH environment. The DWH might simply offer data for decision support to the process, or it may actively influence the control flow of the process.

There are lots of examples showing how important DWHs have become for business processes: When a person applies for a loan in a bank for example, the DWH is an integral part of the loan business process. The applicant is scrutinized to find out if she or he has caused a financial loss previously, or has changed identity and caused damage under a different name. The process of designing a new product in a telecommunication company or an airline, or the composition of the product range in a supermarket requires information on the customer behavior covered by the DWH.

But DWHs can be used for more than just the traditional way of analyzing data. Today, a DWH is an integral part of a highly competitive business environment, where it plays a very active role.

A DWH that is tightly coupled with the operational business, in order to reduce the time between critical business events and the actions taken, is called an active or real-time DWH [1]. Changes in the business process flow can be initiated in near real-time, allowing to move from acting reactively to proactively. For example, when an order cannot be delivered on time, the active DWH of a transportation company can initiate alternative scenarios, such as a different means of delivery or a higher priority for the order. Thus it becomes possible for the company to avoid the delay and the penalty. In this case, the DWH actively initiates changes of the process flow.

Surprisingly, this knowledge – how dynamic business structures interact with the DWH and how the DWH is being used in every day business life – is not made explicit in existing conceptual models. There is a

[1] This research has been funded by the Austrian Federal Ministry for Education, Science, and Culture, and the European Social Fund (ESF) under grant 31.963/46-VII/9/2002.

need for an integrated model of processes and DWHs to make the relationship between the DWH and the business processes more transparent.

To bridge this gap, we extend a business process modeling language with an additional perspective, the *Business Intelligence (BI) Perspective,* to be able to create models that show

- where and how business processes use DWHs,
- which parts of the business processes depend on which parts of the DWH, and
- how the DWH impacts the business process control flows.

A DWH stores decision support data. But the data need to be analyzed and interpreted as well. We therefore adopt the term BI in the name of our perspective, as BI represents a much broader concept. We see BI as all kinds of applications and technologies for storing, analyzing, and providing access to data to help enterprise users to make better business decisions. BI goes beyond DWHs as it covers the entire DWH environment from storage to analysis.

A business process model with a BI Perspective provides the following contributions:

- It shows where business processes depend on or are influenced by BI.
- The BI Perspective can support the design phase of a DWH project, making it possible to prioritize the sub-projects according to business needs.
- It can be used to justify the costs of BI projects by pointing out the unseen relationships between the business processes and its business value, important business decisions, and BI.
- It can be used for estimates of the cost of usage as well as for risk management: If the data quality in a certain area is bad, a data mart fails or data is corrupted, an integrated model enables better reactions because it is known which business processes are affected.
- A process model with the BI Perspective also allows to discover which parts of the DWH are not accessed at all, and decide if those parts should be further maintained.

The business process modeling language we have chosen to extend with the BI Perspective is the Event-Driven Process Chain (EPC). It is introduced in Section 2. The EPC has been chosen because of its wide-spread use and because of its flexible view concept which allows to separate the different aspects of a business process. One can easily add another perspective while keeping the original structure intact.

The BI Perspective is divided into a *Traditional BI Perspective* and an *Active BI Perspective.* It covers the two main types of interaction between the DWH environment and business processes: The *Traditional BI Perspective* (Section 3) shows the different levels and perspectives of a DWH environment which are relevant to a business process. The perspective covers various aspects, from accessing the full DWH or data mart, to accessing single facts and measures, as well as reports.

The *Active BI Perspective* (Section 4) shows how an active DWH influences the control flow of a business process. Related Work is presented in Section 5.

2. Event-Driven Process Chains (EPCs)

The Architecture of Integrated Information System (ARIS) concept [11] involves dividing complex business processes models into separate views, in order to reduce the complexity. The views can be handled independently as well as related. There are three views focusing on functions, data, and the organization (see Figure 1), and an additional view focusing on the integration of the other three.

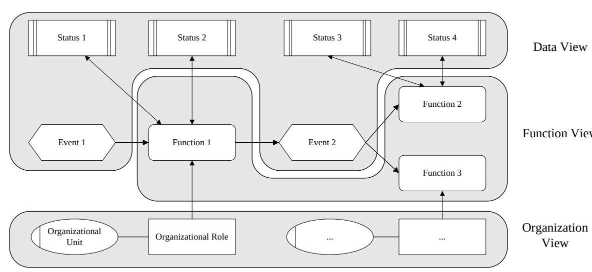

Figure 1: ARIS Views

The *Data View* contains events and statuses. Events such as “customer order received,” or “invoice written” are objects that represent data. Statuses such as “customer status” and “article status” are also represented by data. To provide the data view with a description method for statuses, Chen’s Entity-Relationship (ER) model [4] was adopted into the ARIS framework, since it was the most widespread designing method in the area of data modeling. Today, the UML class diagram is also used [9].

The *Function View* contains the description of the activities to be performed, the individual sub-functions, and relationships that exist between the functions.

The *Organization View* represents the organizational structure. This includes the relationships between organizational units, between employees and organizational units, and employees and their roles.

The *control view* links functions, organization units or roles and data. It integrates the design results, which were initially developed separately for reasons of simplification. The functions, events, information resources, and organization units are connected into a common context by the control flow. The resulting model is the EPC.

The EPC has been developed within the framework of ARIS and is used by many companies for modeling, analyzing, and redesigning business processes. EPCs were developed in 1992 at the Institute for Information Systems of the University of Saarland, Germany, in collaboration with SAP AG. It is the key component of SAP R/3's modeling concepts for business engineering and customizing. The EPC is based on the concepts of stochastic networks and Petri nets.

EPCs are an intuitive graphical business process description language [11]. They describe processes on the level of their business logic, and are targeted to be easy understood and used by business people. The name represents the control flow structure of the process as a chain of events and functions. A basic EPC consists of the following elements (see Figure 2):

- *Functions* are active elements. They model the tasks or activities within the company.
- *Events* are created by processing functions or by actors outside of the model. An event may act as a pre-condition of one or correspond to the post-condition of another function.
- *Logical operators* connect functions and events. There are three types of logical operators: AND, XOR (exclusive or) and OR.

The extended EPC consists of the following elements:

- The *Organization Unit* or *Role* is responsible for performing a specific function.
- The *Information Objects* portray input data serving as the basis for a function, or output data produced by a function. They correspond to entities or attributes of the ER model or classes and attributes of the UML class diagram.

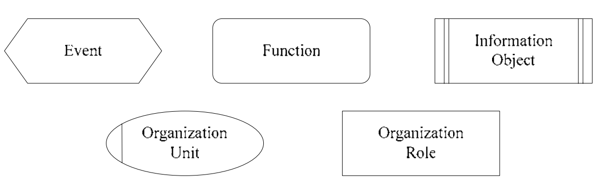

Figure 2: EPC Elements

3. Traditional BI Perspective

We extend the EPC with features for BI modeling to enable the creation of models that integrate information about where the process makes use of decision support data. These models make the hidden knowledge about the relationships between the business processes and BI explicit. We have divided our work into two parts. This section describes the Traditional BI Perspective, which can be used to model scenarios in which business processes access DWHs, data marts or reports followed by a section about the Active BI Perspective, where the DWH actively influences the control flow of business processes.

3.1. The Extended Meta-Model of the EPC

We have chosen the EPC as a basis for our model because of its wide-spread use in many companies for modeling business processes, and because of its flexible view concept, that allows to separate the different aspects of a business process. We can easily add another perspective while keeping the original structure intact.

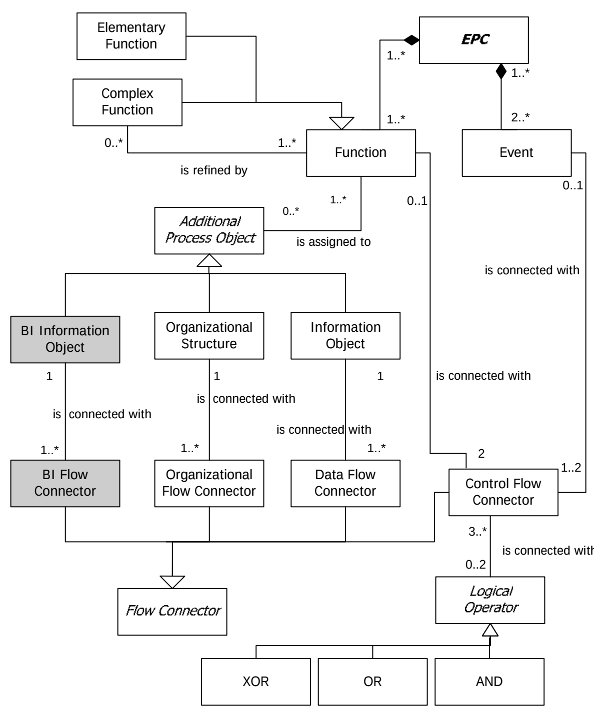

Figure 3: EPC Meta-Model with BI Information Object and BI Flow Connector

The EPC meta-model (white) including the *Traditional BI Perspective* (dark) is shown in Figure 3. Each EPC consists of one or more *Functions* and two or more *Events*, as an EPC starts and ends with an event and requires at least one function for describing a process. A function can be either a *Complex Function* or an *Elementary Function*. Complex functions are refined by one or more other functions. A function is connected with two *Control Flow Connectors*. An event is connected with one or two control flow connectors. Control flows link events with functions, but also events or functions with *Logical Operators*. A logical operator can be either XOR, OR or AND. A logical operator is connected at least with three control flows, one or more incoming and one or more outgoing connectors.

An *Additional Process Object* may be assigned to one or more functions. *Information Objects* and *Organizational Structures* are called additional process objects. All types of additional process objects may be assigned to one function. The organizational structure is connected with one or more *Organizational Flow Connectors*. The information object is connected with one or more *Data Flow Connectors*.

We extend the EPC meta-model with what we call the Traditional BI Perspective and introduce a *BI Information Object* as additional process object. The detailed meta-model of the Traditional BI Perspective is shown in Figure 5. All elements of the model are specializations of the *BI Information Object.* All BI information objects are *Additional Process Objects* in terms of the EPC, which means that they can be assigned to a function that uses the information contained in them. The BI information object is connected with one or more *BI Flow Connectors*, the notation of which is shown in Figure 4.

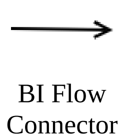

Figure 4: Notation of the BI Flow Connector

What is a BI information object? It represents the ways in which a business process might access decision support data. The business process might see the DWH as a whole, or focus on individual data marts. A subprocess or function can use individual entities and attributes, depending on the level of detail of the process model. With or without DWHs, decision makers often receive relevant data in form of reports, e.g. a report on sales data for the past fiscal year.

We have identified three main categories of BI information objects: *BI Data Repositories* (representing the DWH architecture and what repository is accessed), *BI Data Objects* (representing the data model of a certain repository), and *BI Information Presentation Objects* (representing the means of presentation, either a report or an interactive analysis). They are shown in Figure 5. The objects in each category can be related to each other.

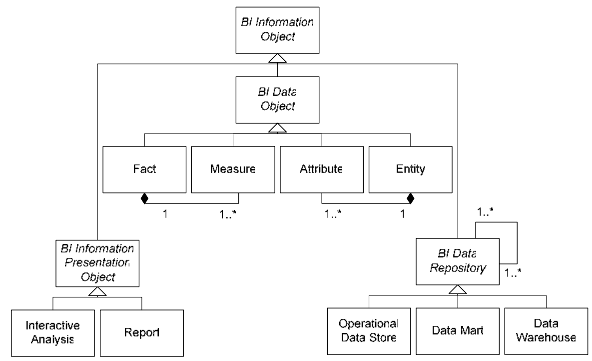

Figure 5: Meta-Model of Traditional BI Perspective

3.2. BI Data Repositories

BI data repositories are the first type of BI information object that can be modeled in relation to a business process. They basically represent different types of databases as used in DWH settings. The types of BI data repositories occurring in a given situation depend on the DWH architecture in an organization. Also, several different data repositories may exist in parallel. Our approach is not limited to any specific DWH architecture, but can be applied to a wide selection of architecture types.

Depending on the architecture, different combinations of BI data repositories may occur in an organization. In large multinational organizations it is not uncommon to have more than one DWH. A large DWH often co-exists within an organization with smaller data marts (DM), departmental subsets of a DWH focused on selected subjects [1]. The data marts might be created based on the data warehouse, obtaining their data from there, meaning that a data mart acts as a kind of materialized view on the DWH. In another situation, the data marts may be created individually by departments without an underlying DWH, and then later be integrated into an organization-wide DWH, making possible operations that span several data marts. Also, there may be none, one or more operational data stores (ODS), located between the operational systems and the DWH [6]. Depending on the architecture, end user applications may query individual data marts and/or the DWH, or even access the data in the ODS directly.

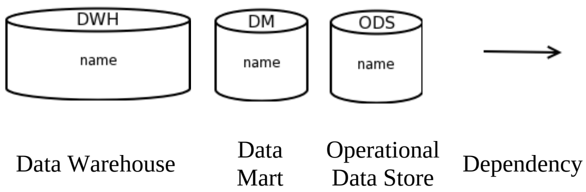

Figure 6: Notation of BI Data Repositories

In order to allow the greatest possible flexibility and provide meaningful content in the models, we have identified three basic types of BI data repositories: the *DWH*, the *Data Mart (DM)* and the *Operational Data Store (ODS)*. They are related to each other through data dependencies. The notation for these elements is shown in Figure 6.

To illustrate the relationships between the BI data repositories we propose a simple repository dependency diagram. In the example diagram shown in Figure 7, the DWH depends on two independent ODS systems, and in turn supplies four data marts with data.

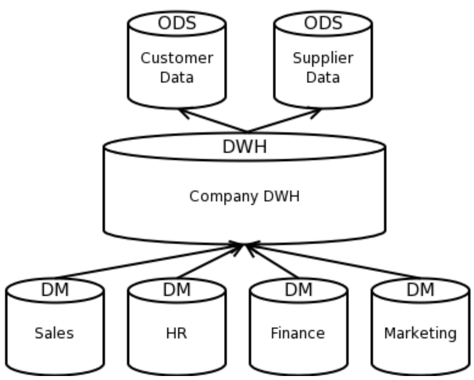

Figure 7: Data Repository Dependency Diagram

3.3. BI Data Objects

In order to provide a more detailed view of the data accessed by the functions of an EPC, we also want to model the individual data entities contained in the BI data repositories. These *BI Data Objects* are generally represented by conceptual data models. For example, if a function needs data on the revenue of a certain product range, it can be modeled to access the corresponding BI data object directly.

Depending on the type of repository, the overall architecture, and the preferences of the designers, different kinds of data models can be used. The two main types relevant to BI applications are entity-relationship modeling and multidimensional modeling [5]. In the first case, we use the Entity-Relationship model [4] and in the latter, the Multidimensional Entity Relationship (ME/R) model [8] described in sub-section 3.4.

The BI data objects of an E/R model that can be accessed by an EPC are either *Entities* or individual *Attributes*. In the case of the ME/R, they are *Facts* or *Measures*. The notation of BI data objects is described in Figure 8. Whether entities / facts or attributes / measures are to be used as additional process objects in the EPC depends on the granularity of the EPC functions.

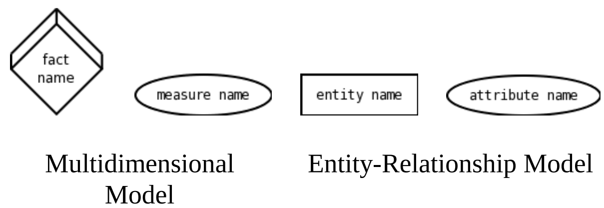

Figure 8: Notation of BI Data Objects

3.4. Multidimensional ER Model (ME/R)

DWH applications involve complex queries on large amounts of data, which are difficult to manage for human analysts. Relational data models "are a disaster for querying because they cannot be understood by users and they cannot be navigated usefully by DBMS software" [8]. In data warehousing, data is often organized according to the multidimensional paradigm, which allows data access in a way that comes more natural to human analysts. The data is located in n-dimensional space, with the dimensions representing the different ways the data can be viewed and sorted (e.g. according to time, store, customer, product, etc.).

A multidimensional model, also called star schema or fact schema, is basically a relational model in the shape of a star. At the center of the star there is the fact table. It contains the subject of analysis (e.g. sales, transactions, repairs, admissions, etc.). The attributes of the fact table (e.g. cost, revenue, amount, duration, etc.) are called measures. The spokes/points of the star represent the dimensions according to which the data will be analyzed. The dimensions can be organized in hierarchies that are useful for aggregating data (e.g. store, city, district, country). Stars can share dimensions, creating a lattice of interconnected schemas that makes drill-across operations possible.

We have chosen the Multidimensional Entity Relationship (ME/R) Model [8] as the conceptual model for a multidimensional data model for our purpose because of its simplicity and expressiveness. ME/R extends the E/R model by adding three elements that are specializations of existing E/R elements.

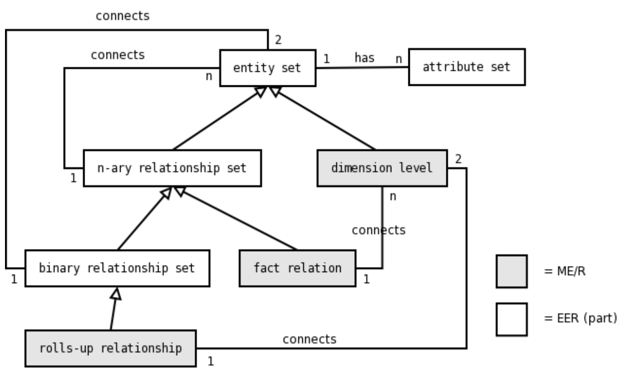

Figure 9: ME/R Meta-Model

The ME/R meta-model is shown in Figure 9. The white elements are part of the E/R meta-model and the gray elements denote the additions by the ME/R. A fact is defined as a *fact relation*, because it is a specialization of the basic *n-ary relationship set*, connecting n dimensions. Because dimensions are organized in hierarchies, they are divided into *dimension levels*. A dimension level is a specialization of an *entity set*, connected with n other dimensions via a fact relation. Within the hierarchy of a dimension, the dimension levels are related to each other via *rolls-up relationships*. Each rolls-up relationship connects two neighboring dimension levels (e.g. product rolls up to category, or region to country). The notation of the ME/R elements is shown in Figure 10.

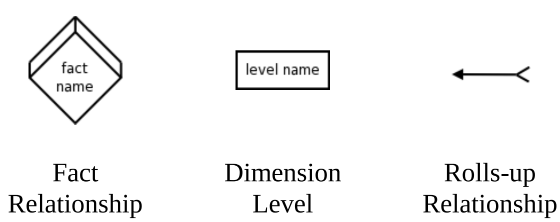

Figure 10: ME/R Notation

The example in Figure 11 shows a simple example of a fact relation, inspired by [8], described in ME/R notation. The *Sales fact* has three dimensions, *Time*, *Product* and *Store*. The levels of the dimensions are only shown for the store dimension. Each entry in the fact table contains information about a single sales event, meaning that one or more items of a product were sold at a store at a given time. For each sale the revenue achieved, the cost incurred, and the quantity sold can be analyzed, as well as aggregations such as "total revenue of a product in all stores in one year". Several such facts can be connected by sharing the same dimensions, creating a more complex multi-cube model.

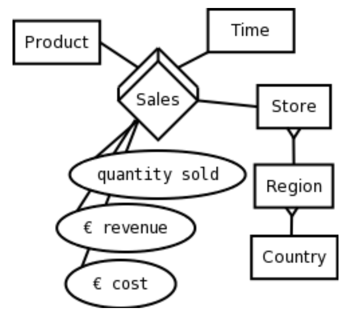

Figure 11: ME/R Notation: Example

3.5. BI Information Presentation Objects

In an organization employing BI techniques, there are usually tools and applications providing users with prepackaged information that has been compiled for them. We call these collections of information BI Information Presentation Objects, and have identified two different types: Report or Interactive Analysis. A report is a kind of document containing the values of measures related to a certain area, for example a report on sales in the south region for the 4th quarter of 2004. The values contained in a report do not change over time. An interactive analysis on the other hand is closer to a tool. It provides its users with regularly updated values and can be used for continuous performance monitoring. In the EPC we can show e.g. a certain report that a function accesses. The notation for BI information presentation objects is shown in Figure 12.

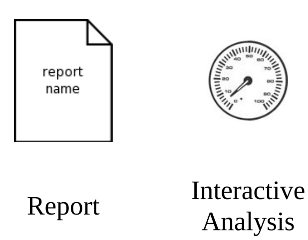

Figure 12: Notation of BI Information Presentation Objects

3.6. EPCs and the Traditional BI Perspective

The BI information objects introduced in the previous sections extend the EPC meta-model as additional process objects. Therefore, they can be used in any EPC diagram. A BI information object connected to a function indicates that the function requires data provided by the BI object.

BI is widely used in the area of fraud detection. The example EPC in Figure 13 demonstrates the application of the Traditional BI Perspective in such a case. Any company selling goods online is faced with the problem of credit card fraud. In this example freud detection business process, every order is first subjected to a number of automated checks. Depending on the result of the first steps, only suspicious cases are investigated further, in order to reduce the workload of the fraud detection department and reduce overall costs.

The automated checks preformed for each order include, among other things, methods such as checks against lists of known offenders, identity verification of the people involved, plausibility checks, and the authorization status of the credit card.

The function *Automated Checks* is started by the event of an order arriving, and does not need access to the DWH. As a result of this function, the claim is either *Red* (client has been involved in previous problems), *Yellow* (suspicious) or *Green* (not suspicious).

For any order classified as *Red*, the customer is asked to phone the customer service department. All claims flagged as *Yellow* are forwarded to the fraud detection department for review and further analysis. In this step, the details of the data used in the previous step are reviewed and compared to details and history of similar cases. The reviewer then decides whether the risk of potential fraud is enough to justify a formal investigation of the case. False positives generated by the automated checks are likely to be identified in this step. They are re-classified as *Green* and will be processed normally.

The function *Review* is performed by the fraud detection department. The function accesses the information contained in the facts *Customers* and *Credit Card Transactions*. The output events of

this function indicate that the claim has either been identified as *Potential Fraud* or as *Green*.

The claims likely to be fraudulent are analyzed in depth by a fraud detection specialist. In order to obtain certainty concerning the final assessment of the claim, the investigator can explore the entire data warehouse of the company, searching for information that answers the remaining open questions. This includes access to external data sources. If the order has been successfully identified as fraudulent, it is flagged red and denied. Again, if the investigation reveals that the order is genuine, it is processed normally.

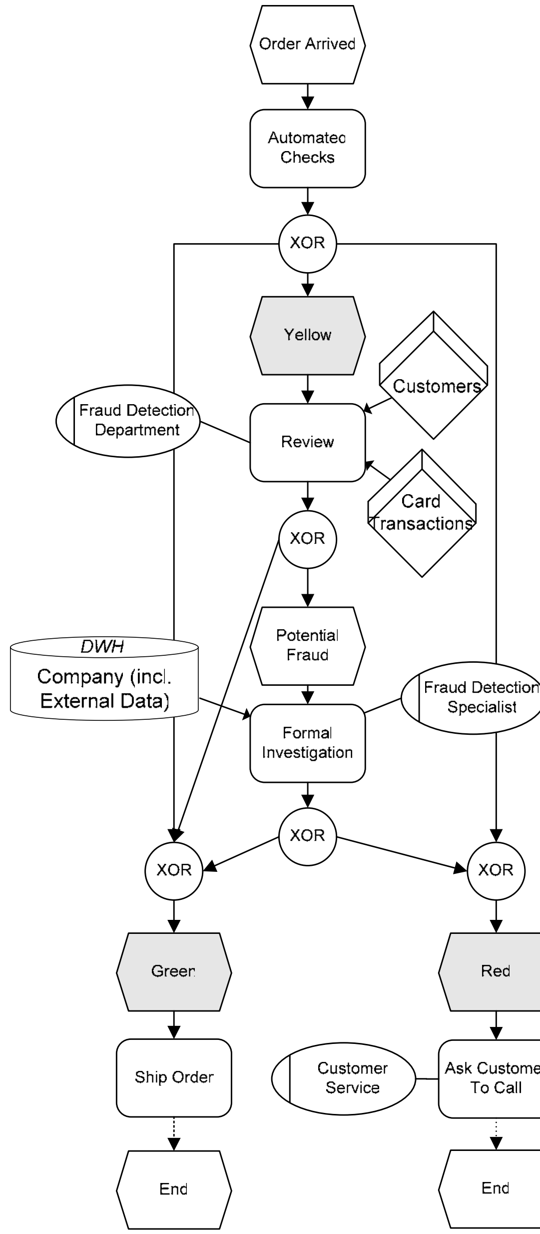

Figure 13: Example EPC with BI Elements

In the EPC, this step is modeled by the function *Formal Investigation*. It needs access to the whole data warehouse, which, in addition to the usual wealth of company internal data, also contains external data. This is typical of a process step that cannot be predefined in detail. It strongly depends on the individual situation, which information can be of use to the investigating specialist. Again, the function is followed by an XOR-operator, either leading to the event that the order is classified as definitely red or green.

3.7. Macro Level

Detailed diagrams such as the one in Figure 13 provide a high level of detail on the flow of a business process, which is not necessary for all purposes. On a macro level, business processes are modeled as black boxes whose internal complex process logic is hidden for sake of clarity. Such diagrams are useful for providing a large scale overview, e.g. over the way a company does business, using a process landscape diagram showing all processes and their interactions with other processes or customers or partners. Diagrams at a macro level give a comprehensive understanding and highlight the relationships with, or the dependencies on other objects.

Such a macro view can also be combined with BI information objects. In order to show the dependencies between BI information objects and business processes we introduce a business process dependency diagram. This diagram may show several different types of cases: Firstly, a business process depending on one or several BI data repositories. This shows the variety of BI data repositories supplying a business process with data. If only one BI data repository fails, the entire business process will be affected and may even stop operating. Secondly, business processes depending on a BI data repository. This shows where the data from a certain BI data repository are used, and which business processes are affected when a BI data repository fails, or provides bad data quality. In both cases, instead of BI data repositories, reports can also be used. Figure 14 shows a business process dependency diagram that illustrates on which data marts a business process relies on. The process *Product Development* employs information provided by three data marts: *CRM*, *Sales*, and *Marketing*.

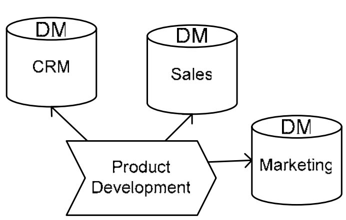

Figure 14: Business Process Dependency Diagram with BI Data Repositories

From another point of view, we can use the business process dependency diagram to show where the data from a certain data mart is used. Figure 15 provides an example showing the *Sales* data mart providing the business processes *Controlling*, *Product Development* and *Management Strategy* with data.

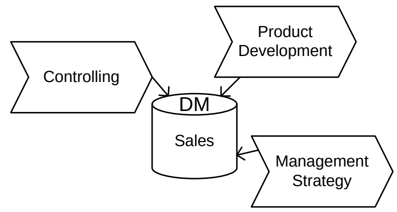

Figure 15: Business Process Dependency Diagram with BI Data Repositories

On the macro view BI presentation objects can also be combined with business processes. A Report supplies business processes with data and a business process depends on data from one or more reports. In order to illustrate these relationships, a business process dependency diagram with reports as shown in Figure 16 can be used. This shows the variety of reports supplying a business process with data. If only a single report contains invalid data, the whole process is affected.

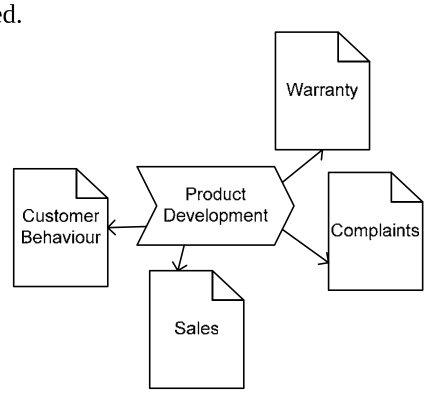

Figure 16: Business Process Dependency Diagram with Reports

4. Active BI Perspective

In a process-driven enterprise, BI should not only be seen as a pure static retrieval of preprocessed data to support decision makers. With the control of business processes in near real-time through knowledge and awareness of current business situations, BI moves from acting reactively to proactively. An essential prerequisite for this step is a process-oriented embodiment of BI, thereby becoming an integral part of the business process. Active BI is a dynamic discovery process which continuously:

- Observes and collects events from a business environment.
- Converts the event data into meaningful business information.
- Discovers and analyzes business situations and exceptions.
- Automatically decides the most appropriate actions for a response to the business environment.
- Executes the business actions based on the decision that has been made.

This response can either change the state of a business process or notify humans or IT systems that may be interested in the outcome and result of the decision making.

In this section we extend the EPC with elements that enable the creation of models that integrate information about where the control flow of a business process is changed by BI data.

4.1. Characteristics of BI Processes

BI processes generate new knowledge and business information for supporting business processes. The main objectives of BI processes are to:

- Discover situations and exceptions in business processes. For instance, organizations want to detect suspicious customer behavior (e.g. fraud behavior) which can be countered with a proactive response.
- Provide proactive responses by continuously observing and analyzing customers, business partners and the competition, business processes can be proactively adapted and optimized (e.g. continuous update and optimization of production plans based on the current orders of customers).
- Analyze business data in real-time, in order to change processes instances during execution.

BI processes represent what we call a *Sense and Respond Loop*, which can be divided into 5 stages. In Table 1, the phases are described.

Table 1: Stages of Sense and Respond Loops

Stage	Scope	Description
Sense	What is the current state of the business environment?	Events are unified, cleaned and prepared before the actual analytical processing begins.
Interpret	What do the captured events indicate? What do the event data mean for the current situation of the organization?	Transformation of the events into business information, such as key performance indicators. Discovery of business situations and exceptions.
Analyze	Which business opportunities and risks can arise? What are the possibilities to improve the current situation of the organization?	Analysis of key performance indicators and determination of root causes for business situations and exceptions. Prediction of the performance and assessment of the risks for changing the business environment.
Decide	Which strategy is the best to improve the current situation of the organization? What are the actions required to successfully put a decision into action?	Selection of the best option for improving the current business situations and determines the most appropriate action for a response to the business environment. This step can be automated with rules or by involving humans (a decision maker selects from alternative choices; the Sense and Respond loop continues the processing with the choice of the decision maker).
Respond	Who has to implement the decision? How can the decision be put into action?	Response to the business environment by communicating the outcome to the business process with events.

4.2. The Extended EPC Meta-Model

In the following we extend the EPC meta-model (Figure 3) with *BI Process Objects* (grey elements in Figure 17) that link EPCs with BI processes. Through *Event Flow Connectors*, a *Sense Delegate* receives and forwards business process events, which start the BI process. The *Respond Delegate* communicates the outcome of the sense and respond loop, i.e. the BI process.

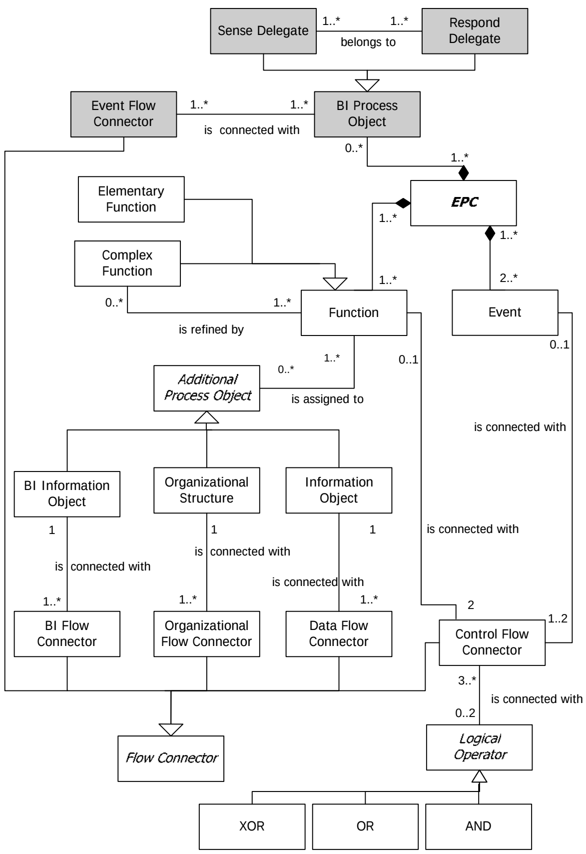

Figure 17: EPC Meta-Model with BI Process Objects and Event Flow Connector

Figure 18 shows an EPC which is supported by a BI process. Events of the EPC are linked with a sense delegate, which represents the collection and forwarding of these events which become the input for the BI process. Therefore, the sense delegate is a sensor that emits the business process events to the corresponding BI processes. On the other hand, respond delegates communicate an event about the outcome of the BI process.

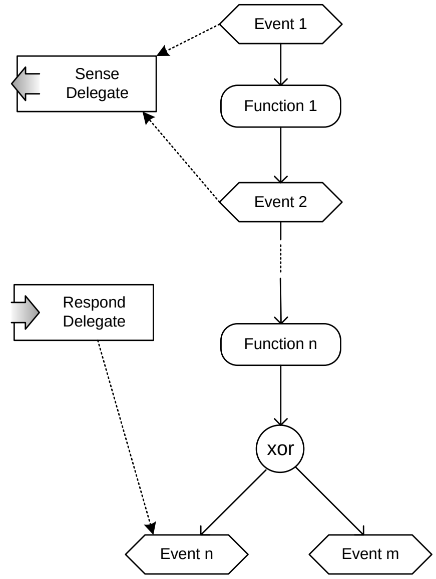

Figure 18: EPC with a BI Process

4.3. The Meta-Model of the BI Process

The objectives listed in section 4.1. indicate that BI processes are very event-driven in their nature and focus on processing in order to deliver results in near real-time. BI processes generally do not require much interaction with humans and external systems of business partners. Furthermore, processing steps depend mostly on the availability of processing results (e.g. analytical results, discovered situations) from other steps. Therefore, in a BI process, event flows have a higher importance than control flows.

For these reasons, we decided to propose a new model that is tailored to the requirements of BI processes. EPCs do not show direct data flows between functions which makes it difficult to model data-driven tasks. The BI meta-model (Figure 19) includes various elements which can be used to construct a BI process. External interfaces to business processes are represented by *Sense Delegates* and *Respond Delegates*. A *Processing Step* represents a unit of work that has to be performed within the BI processes (e.g. the calculation of a performance indicator or a data analysis). *Event Flow Connectors* represent event streams and connect the processing steps within the BI process. They show how events flow from one processing step to the next. *Connection Points* facilitate merging and splitting event streams.

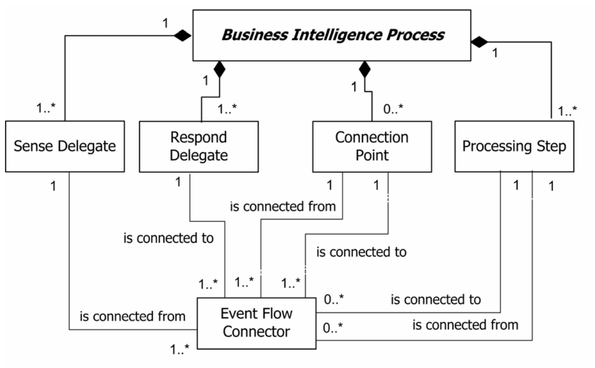

Figure 19: Meta-Model for BI Process

Since the connectors between delegates and processing steps are event flows (which are essentially data flows), our model does not use any typical elements for control flows (e.g. XOR or OR elements in EPCs) to reflect that BI processes are generally data-driven in their nature. An event flow essentially represents a data flow which conforms to a certain event type. The notation of the BI process elements is shown in Figure 20.

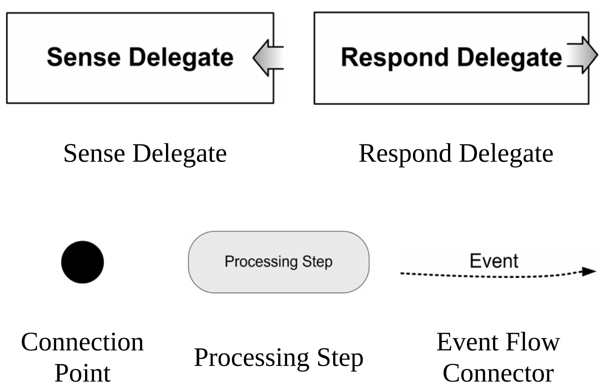

Figure 20: Notation of the BI Process Elements

Please note that the sense and respond delegates in a BI process do not have to originate from the same business process. BI processes can support multiple business processes. For instance, if a BI process is used to intelligently control the production process based on the currently placed orders, the sense delegate would be part of the order process and the respond delegate would be part of the production process. In other words, with our model BI processes can be used to intelligently adjust business process behavior across the entire enterprise.

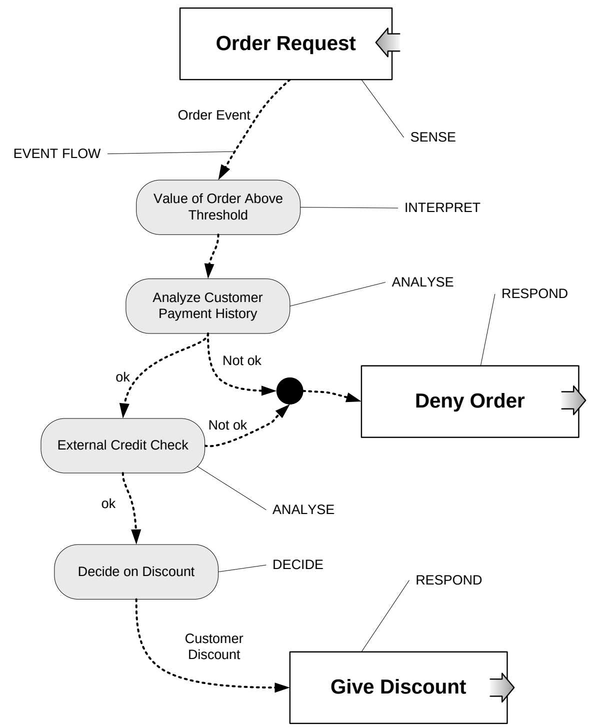

Figure 21: BI Process

In the example shown in Figure 21, the sense delegate *Order Request* receives order events occurring in the business process. Each event passes through a number of processing steps. In the first step, all order events are interpreted. Only if an order with a value higher than a certain threshold is identified, it is passed on to an analyzing step. In the case of this example the payment history of the customer involved is reviewed. Depending on the outcome of the analysis, the order request is either denied, e.g. the event is passed on to the *Deny Order* respond delegate, or a second analysis step, an External Credit Check, is performed. Again, depending on the outcome of the check, the order may be denied. In the case of a satisfactory outcome, a decision on a discount is made. When the amount of the discount has been decided, the *Give Discount* respond delegate responds to the business process.

5. Related Work

There are a lot of conceptual models available for business processes, data bases or DWHs. But there are no models available that focus on the relationship between the DWH and the business processes. EPCs [11] incorporate a data view targeting operational data bases. EPC functions perform read or write operations on the databases and their entities (different to the DWH). Our BI Perspective embodies a similar concept but goes much further.

In UML 2 Activity Diagrams [9] data is represented by data store nodes. A UML Action can perform read or write operations, like the EPC function. The data store node is not necessarily linked with a UML class or database.

The Business Process Modeling Notation (BPMN) [2] provides data objects, which are used and updated during the process. The data object can be used to represent many different types of object, both electronic or physical.

6. Conclusion

In this work we address the missing link in conceptual modeling between business intelligence and business processes. We extended a business process modeling language with an additional perspective, the Business Intelligence (BI) Perspective, to bridge this gap. With the BI Perspective, it becomes possible to create models that show where and how business processes use decision support data, as well as which parts of the processes depend on which information from the DWH environment, and how an active DWH may impact on the business process control flow. The business process modeling language we have chosen to extend with the BI Perspective is the Event-Driven Process Chain (EPC).

The BI Perspecitve in divided into the Traditional BI Perspective and the Active BI Perspective, the latter addressing the specific features of active DWHs. In the Traditional BI Perspective, elements representing the different types of data repositories that are accessed, as well as representing the data model of a certain repository, and additionally elements representing the means of presentation have been designed and incorporated into a meta-model. The Active BI Perspective adds to EPCs the concept of external BI processes that may influence the control flow of business processes. Both parts of the BI perspective have been tested with example business processes.

References

[1] S. Brobst and J. Rarey, “The Five Stages of an Active Data Warehouse Evolution”, *Teradata Magazine* 2001.

[2] Business Process Modeling Notation (BPMN), BPMN Specification 1.0, May 3, 2004, http://www.bpmn.org

[3] S. Chaudhuri and U. Dayal, “An overview of data warehousing and OLAP technology”, *ACM SIGMOD Record*, vol. 26, issue 1, ACM Press, 1997, pp 65-74.

[4] P. Chen, “The Entity-Relationship Model - Toward a Unified View of Data”, *ACM Transactions on Database Systems*, vol. 1, issue 1, 1976, pp 9-36.

[5] M. Golfarelli, D. Maio and S. Rizzi, “Conceptual Design of Data Warehouses from E/R Schema”, *Proceedings of the Thirty-First Annual Hawaii International Conference on System Sciences*, IEEE Press, 1998.

[6] M. Golfarelli, S. Rizzi and I. Cella, “Beyond data warehousing: What's next in business intelligence?”, *Proceedings of the 7th International Workshop on Data Warehousing and OLAP*, ACM Press, 2004.

[7] W. H. Inmon and R. D. Hackethorn, *Using the Data Warehouse*, John Wiley & Sons, 1994.

[8] R. Kimball and M. Ross, *The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling*, John Wiley & Sons, 2002.

[9] Object Management Group, Inc., UML 2.0 Superstructure http://www.omg.org/cgi-bin/doc?ptc/2004-10-02 (18/3/2005).

[10] C. Sapia, M. Blaschka, G. Hoefling and B. Dinter, “Extending the E/R Model for the Multidimensional Paradigm”, *Proceedings of the ER '98 Workshop on Data Warehousing and Data Mining: Advances in Database Technologies*, Springer Verlag, LNCS 1552, 1998, pp. 105-116.

[11] A.-W. Scheer, *ARIS – Business Process Modeling*, Springer Verlag, 1999.